

# WHAT IS EFFECTIVE SCHOOLING?

## A review of current thought and practice

Prof. Jaap Scheerens  
University of Twente, the Netherlands  
[J.Scheerens@utwente.nl](mailto:J.Scheerens@utwente.nl)

*This research was supported through a grant by the International Baccalaureate Organization. The purpose of this research is to collect and synthesize information, including external expert opinion, policy documents and scholarly work, which discuss theories and practices related to school effectiveness. The views expressed in this paper are those of the author and cited work and do not necessarily represent or reflect the views of the International Baccalaureate Organization.*

## Table of contents

Title page.....	1
Chapter 1: Definition and conceptualization .....	4
Chapter 2: Trends in research and policy .....	10
Chapter 3: Best practices in making schools more effective .....	23
References.....	37
ANNEX 1: How the results for table 4 were computed .....	44
ANNEX 2: Further results of meta-analyses .....	44
ANNEX 3: More details on results from international comparative studies.....	47
ANNEX 4: Component and sub-items of fourteen effectiveness-enhancing factors (cited from Scheerens et al., 2007) .....	51
ANNEX 5: Annotated bibliography of key publications.....	73

## Preface

In this report conceptual issues of school effectiveness are discussed and a state of the art review of the knowledge base is presented. In addition, implications for educational policy and practice are discussed.

Key words are: a multi level representation of educational effectiveness, syntheses of reviews and meta-analysis, context dependency as well as generalizability of school effectiveness research findings, international comparative outcomes.

## Chapter 1: Definition and conceptualization

### Preface

In this chapter the term school effectiveness is defined. It is compared to the broader concept of educational effectiveness, compared to school improvement and described as a specific facet of educational quality. Foundational issues, concerning the integrity of the concept across contexts, are also discussed.

### General definition

In the most general sense, 'school effectiveness' refers to the level of goal attainment of a school. Although average achievement scores in core subjects, established at the end of a fixed program are the most probable 'school effects', alternative criteria like the responsiveness of the school to the community and the satisfaction of the teachers may also be considered.

Assessment of school effects occurs in various types of applied contexts, like the evaluation of school improvement programs or comparing schools for accountability purposes, by governments, municipalities or individual schools.

School effectiveness research attempts to deal with the causal aspects inherent in the effectiveness concept by means of scientific methods. Not only is assessment of school effects considered, but particularly the attribution of differences in school effects to malleable conditions. Usually, school effects are assessed in a comparative way, e.g. by comparing average achievement scores between schools. In order to determine the 'net' effect of malleable conditions, like the use of different teaching methods or a particular form of school management, achievement measures have to be adjusted for intake differences between schools. For this purpose student background

characteristics like socioeconomic status, general scholastic aptitude or initial achievement in a subject are used as control variables. This type of statistical adjustment in research studies has an applied parallel in striving for 'fair comparisons' between schools, known under the label of 'value-added'.

### Demarcation between school effectiveness and educational effectiveness

When educational systems are seen as hierarchies, school effectiveness can be distinguished from instructional effectiveness, which plays out at classroom level, and from "system effectiveness". The latter term is less common, and refers to a more recent strand of research that is strongly stimulated by the upsurge of international assessment studies. In such studies policy amenable conditions at the national system level can be associated with student outcomes; examples are policies of enhancing school autonomy, accountability and choice. When school effectiveness depends on school level malleable conditions, instructional (or teaching) effectiveness on activities of teachers, and system effectiveness on policy amenable conditions at the national level, the term educational effectiveness can be used as referring to the union of these three.

At the technical level multi-level analysis has contributed significantly to the development of integrated school effectiveness models. In contributions to the conceptual modeling of school effectiveness, schools became depicted as a set of 'nested layers' (Purkey and Smith, 1983), where the central assumption is that higher organizational levels facilitate effectiveness enhancing conditions at lower levels (Scheerens and Creemers, 1989). Although the focus of this report is on school effectiveness, it is considered more interesting and policy relevant to see school level factors in relation to system level and classroom level variables. This approach could either be described as confirming to the conceptual modeling of integrated school

effectiveness models, or as treating school effectiveness as embedded in educational effectiveness.

## Demarcation between school effectiveness and school improvement

The concept of school improvement may refer to a product (improved performance of a school over time), or to a controlled or emerging process of change that evolves in time, involving procedural aspects and specific content.

When school effectiveness is seen as a research activity; school improvement could be taken as the dynamic application of the research results, i.e. the active manipulation of the “process” conditions identified as correlates of educational outcomes. A first and basic view of linking improvement and effectiveness would therefore be to say that the results of school effectiveness research provide likely content for school improvement. When school improvement is seen as a systematic activity, two extra emphases are usually at stake; firstly that the process of setting in motion effectiveness enhancing conditions is studied as a change process, and secondly that the control of the change process is seen as distinct from routine control of the organization. This means that school improvement goes beyond the direct management of the primary process of teaching and learning but often includes adaptations of the management approach and organizational conditions as well.

The growing interest in both fields (educational effectiveness and school improvement) in longitudinal designs, often referred to as a more dynamic approach, narrows the distinction between them, and makes a complete conceptual integration more feasible (Creemers and Kyriakides, 2012). The role of school management and leadership in school improvement is particularly interesting. In some conceptual models (e.g. Hallinger & Heck, 2010), improvement is the result of school leadership efforts, while changes in leadership approach might also be seen as

part of a school improvement program. External “change agents” may be involved in the latter case.

## School effectiveness as a facet of school quality

A basic system model to depict the functioning of educational systems and schools as organizations is a good analytical tool to define facets of quality that are amenable to empirical analysis and verification. According to this model the school is seen as a black box, within which processes or ‘throughput’ take place to transform inputs into outputs. The inclusion of an environmental or context dimension completes the model (see Fig. 1).

When the level of outputs is the core of quality judgments on schools, educational programs, or the functioning of national educational systems, this could be described as the productivity perspective. There are many practical applications of this perspective: test based accountability policies, school performance feedback systems, and the comparison of mean country level achievement among countries, on the basis of international assessment studies, like TIMSS and PISA. In case the interest is not focused primarily on average achievement levels, but rather on the distribution of outcomes, inputs and processes, equity is the predominant quality facet. In international comparisons equity is getting more and more attention (see for example the OECD report titled “Overcoming social background”, based on the 2009 edition of PISA (OECD, 2010)). At the school level Inspection Frameworks may contain indicators on equity (Janssens, 2007). When effectiveness is the predominant quality perspective, the focus is on the instrumental value of input and process indicators to maximize output. This is the question on “what works best”. From a quality perspective this means that it is not the “beauty” of organizational arrangements or teaching strategies, but the extra value these approaches create in terms of school output. In a subsequent chapter prac-

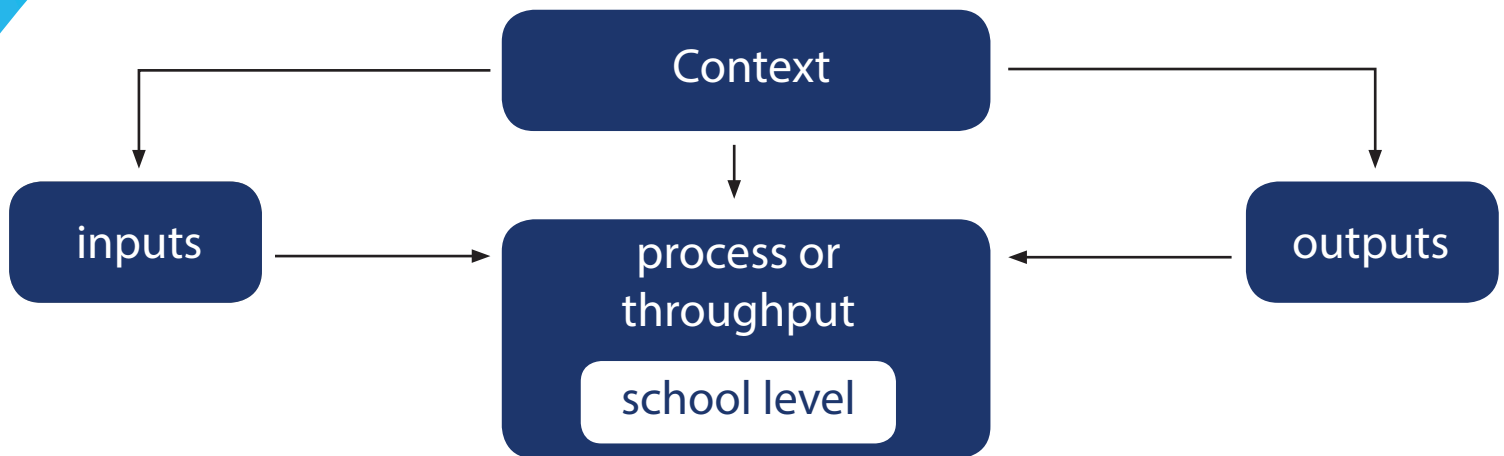


Figure 1: A basic systems model of school functioning

tical implications of the effectiveness perspective will be discussed in more detail. When effectiveness at the lowest possible costs is considered efficiency is the quality facet in question. Monetary measures of inputs are key aspects in efficiency measurements. Finally, the relationship of the school with its environment or context may be the core issue for quality judgments; particularly the question of responsiveness, which in the most general sense means that a school pays attention to impulses, both in terms of supply and demand, from the larger context. Where effectiveness and efficiency deal with the question of “doing things right”, responsiveness may be seen to address the question of “doing the right things”, such as choosing educational objectives that confirm to the demands of further education or the labor market.

These facets of educational quality, defined on the basis of their key elements and interrelationships included in Figure 1, are schematically summarized here:

Quality facet	Key indicators and relationship between indicators
Productivity	outcomes
Equity	The distribution of inputs, processes and outcomes
Effectiveness	Association between inputs and processes on the one hand and outcomes on the other
Efficiency	Effectiveness at the lowest possible costs
Responsiveness	The way input, processes and intended outcomes are fitted to the demands of the context

Two final remarks with respect to effectiveness as a facet of school quality are in order. Firstly, it should be noted that effectiveness refers to causality between means and ends in a complex practical situation, and therefore is analytically difficult. Secondly, this very characteristic of being centered on malleable “causes” of intended effects also points

at great practical relevance, namely its potential for school improvement.

## Strands in educational effectiveness research

Research tradition in educational effectiveness varies according to the emphasis that is put on the various antecedent conditions of educational outputs. These traditions also have a disciplinary basis. The common denominator of the six areas of effectiveness research that will be distinguished is that in each case the elementary design of associating outputs or outcomes of schooling with policy amenable conditions (inputs, processes or contextual) applies.

The following research areas or research traditions can be distinguished:

- 1) Research on equality of opportunities in education and the significance of the school in this.
- 2) Economic studies on education production functions.
- 3) The evaluation of compensatory programs and school improvement programs.
- 4) Studies of unusually effective schools.
- 5) Studies on the effectiveness of teachers, classes and instructional procedures.
- 6) Studies on the effectiveness of system level policies and institutional arrangements

For a further discussion of the first five of these research traditions the reader is referred to Scheerens, Glas and Thomas (2003, Ch. 11). A schematic characterization of research orientation and disciplinary background is given in Table 1. The 6th research area is an emerging field, which is very much stimulated by international assessment programs, such as TIMSS and PISA. This is the case, because only international comparative studies allow for the analyses of the way country level characteristics of educational system vary between countries. System level variables that have been addressed in this kind of study are decentralization, choice and accountability arrangements in national

	<b>independent variable type</b>	<b>dependent variable type</b>	<b>discipline</b>	<b>main study type<sup>1</sup></b>
a. (un)equal opportunities	socio-economic status and IQ of pupil, material school characteristics	attainment	sociology	survey
b. production functions	material school characteristics	achievement level	economics	survey
c. evaluation compensatory programs	specific curricula	achievement level	interdisciplinary pedagogy	quasi-experiment
d. effective schools	'process' characteristics of schools	achievement level	interdisciplinary pedagogy	case-study
e. effective instruction	characteristics of teachers, instruction, class organization	achievement level	educational psychology	experiment observation
f. system level effectiveness	system level policies and institutional arrangements	achievement and attainment	economics	background studies based on international assessment programs

*Table 1: General characteristics of types of school effectiveness research*

<sup>1</sup> It should be noted that currently the more basic study types are frequently blended in more comprehensive data systems

educational systems (cf Woessmann, et al., 2009). Scheerens, et al. (2013) in a study that used PISA 2009 data, investigated more complex models in which indirect effects of system level factors through intermediary school conditions on student performance were computed.

A very interesting methodological development is the new interest in the use of randomized field trials in school effectiveness research (Bosker, 2011). Many relevant examples are presented on the website attached to the bi-annual SREE (Society for Research on Educational Effectiveness) conferences (<https://www.sree.org/>).

When integrated models of school effectiveness were introduced, in which the above strands were combined, the multi level nature of such models was also emphasized. Next, integration also implied an interdisciplinary orientation. As a matter of fact a synthesis between production functions, instructional effectiveness and school effectiveness became possible, by including the key variables from each tradition, each at the appropriate 'layer' or level of school functioning [the school environment, the level of school organization and management, the classroom level and the level of the individual student]. Conceptual models that were developed according to this integrative perspective are those by Scheerens (1990), Creemers (1994), Stringfield and Slavin (1992), and Creemers and Kyriakides (2008). By way of illustration the Scheerens model is shown in Figure 2.

Exemplary cases of integrative, multi-level school effectiveness studies are those by Mortimore, et al. (1988), Hill, et al. (1996), Sammons,

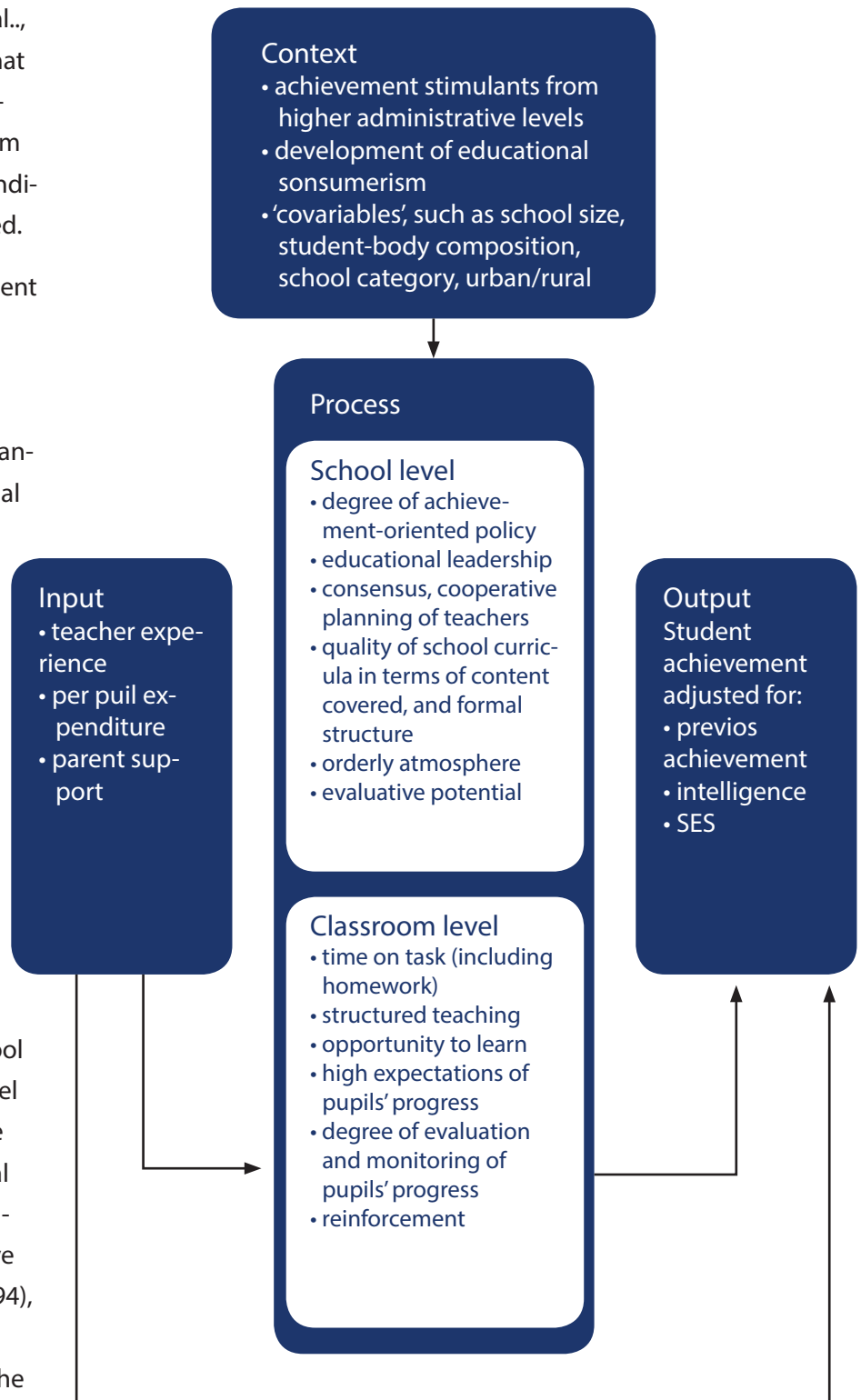


Figure 2: An integrated model of school effectiveness (from Scheerens, 1990)



et al. (1995) and Grisay (1996). The study by Borman, et al (2003) provides a review and meta-analysis of evaluations of Comprehensive School Reform (CSR) projects. CSR projects form a very interesting blending of school effectiveness and school improvement, as well as an application of integrating effectiveness enhancing conditions at school context, school and classroom level.

## Foundational issues in school effectiveness research

The question “what are genuine school effects” addresses the conceptual integrity of the concept of school effectiveness. In school effectiveness research we are interested in the magnitude of the effect of going to one school as compared to the next, and to the degree this effect can be explained by malleable conditions defined at the school level.<sup>2</sup> With respect to the first question we would speak of a genuine school effect when this effect would be the same, regardless of whether it would be assessed at a certain grade level, for a certain school subject and in a particular year. Consistency in the estimation of school effects across grades, teachers and subjects, and stability of school effects across years can be seen as foundational issues in school effectiveness research. Several authors have addressed this issue by means of analysis of a correlation matrix of subject- and cohort (or grade) level effects, and computing the magnitude of a general school factor. Bosker, (1990) found a school factor in secondary schools in the Netherlands, that accounted for 70% of the (gross) subject and cohort specific school effects. Van der Werf and Guldmond (1995) carried out the same kind of analyses, based on value-added school effects in primary schools (subjects: arithmetic and language), and found a common school factor that explained 39% of the total

between cohort and between subjects effect variance. A similar decomposition was carried out by Luyten, (1994) using secondary school data. Luyten analyzed gross school effects (unadjusted for initial achievement or socio economic background), studying five cohorts and 17 subject domains. He found a consistent and stable school effect (across subjects and years) of only 25%. In his study the subject effect was 40%, the year effect 8% and the year/subject interaction 27%. In organizational terms the subject effect coincides with the departmental structure of secondary schools in the Netherlands, which in this study was stronger than the school effect. These results draw attention to internal segmentation of schools as organizations, and point at likely overestimation of school effects, when variation between subject matter domains, grades and teachers are not taken into consideration. As such these results underline the importance of integrated school effectiveness models, and multi-level analyses.

The stability of school effects is an issue that becomes practically relevant in situations where schools are compared for their excellence, as part of accountability and/or incentive schemes. Typically the rank ordering of the (value-added) mean achievement of schools is correlated across years. Bosker et al. (1989) found correlations that declined according to the time interval from one to four years from .74 (one year), .62 (two years), .49 (three years) and .49 (four years) in a study of Dutch secondary schools. Gray et al. (1995), looked at time intervals of one, two and three years in English secondary schools and found correlations of .94, .96 and .81. Thomas et al. (2010) analyzed school data over a period of 11 years in the Lancashire district. They concluded that there was a fair stability in school effects. Still, when schools were categorized as average, over- or underachieving there were many switches,

---

<sup>2</sup> It should be noted that school effectiveness has been frequently studied on the basis of naturally occurring variation in school performance; a second approach has it connected to interventions, and research designs that resemble program evaluations or (quasi) experiments, in which effectiveness can also be judged against pre set norms or criteria. The work on stability that is cited in the text is based on the first approach, where developments over time in a sample of schools are studied.

and over a period of 11 years 50% of the schools had changed category. Moreover continuous progress was rare:

*"For the majority of schools three years of upward movement seems to have been the typical limit. In short, our evidence from the non-linear modelling suggests that, whilst there were undoubtedly changes, these were not very 'continuous' and in many cases could have occurred by chance. This finding contrasts starkly to government ideals of continuous school improvement." (Thomas et al., 2010, p. 280)*

Less stability was again found in a recent Dutch study, where it appeared that of the highest scoring secondary schools, only 15% were still in the top category three years afterwards, Vermeer and Van der Steeg (2011).

These results show that the stability of school effects may vary across countries. As a caution against instability it would make sense to assess the position of schools in accountability and reward schemes over a certain period of time, say three years; and compare schools on their average achievement across these three years.

## Chapter 2: Trends in research and policy

### Preface

The first part of this chapter discusses the school variables that are most commonly addressed in school effectiveness research as well as their effect sizes, in terms of association with student achievement. It appears that there is conformity on the former (selection of variables) but less consensus on the effect sizes. Specific attention is given

to the research results that are based on internationally comparative assessment studies. Results from these studies show generally lower effect sizes than research studies within countries.

The second part of the chapter analyses policy measures to enhance school effectiveness. A multi level conceptual framework, including system, classroom and student level variables, next to school variables, is presented to contextualize improvement levers. Specific attention is given to system level policy amenable variables. Next various approaches to school improvement are described: performance oriented systemic reform, school based improvement inspired by a social psychological orientation and Comprehensive School Reform.

### I. Research

#### Identification of effectiveness enhancing school conditions; consensus among reviews

The core of educational effectiveness research is the identification of effectiveness and improvement oriented conditions. In this section recent and earlier research reviews will be cited, and considered for consensus on the main effectiveness enhancing conditions.

Table 2 summarizes the characteristics listed in somewhat older reviews by Purkey and Smith (1983), Scheerens (1992), Levine and Lezotte (1990), Sammons et al. (1995), Cotton (1995).

Purkey & Smith, 1983	Levine & Lezotte, 1990	Scheerens, 1992	Cotton, 1995	Sammons, Hillman & Mortimore, 1995
Achievement-oriented policy	Productive climate and culture	Pressure to achieve	Planning and learning goals	Shared vision and goals
Cooperative atmosphere, orderly climate		Consensus, cooperative planning, orderly atmosphere	Curriculum planning and development	A learning environment, positive reinforcement
Clear goals on basic skills	Focus on central learning skills		Planning and learning goals school wide emphasis on learning	Concentration on teaching and learning
Frequent evaluation	Appropriate monitoring	Evaluative potential of the school, monitoring of pupils' progress	Assessment (district, school, classroom level)	Monitoring progress
In-service training/staff development	Practice-oriented staff development		Professional development collegial learning	A learning organization
Strong leadership	Outstanding leadership	Educational leadership	School management and organization, leadership and school improvement, leadership and planning	Professional leadership
	Salient parent involvement	Parent support	Parent community involvement	Home school partnership
Time on task, reinforcement, streaming	Effective instructional arrangements	Structured, teaching, effective learning time, opportunity to learn	Classroom management and organization, instruction	Purposeful teaching
High expectations	High expectations		Teacher student interactions	High expectations
				Pupil rights and responsibilities
			Distinct-school interactions	
			Equity	
			Special programs	
		External stimuli to make schools effective		
		Physical and material school characteristics		
		Teacher experience		
		School context characteristics		

Table 2: Effectiveness enhancing conditions of schooling in five early review studies (*italics in the column of the Cotton study refers to sub-categories*). Source: Scheerens, 2000

Consensus is largest with respect to the factors:

- achievement orientation (which is closely related to “high expectations”);
- co-operation;
- educational leadership;
- frequent monitoring;
- time, opportunity to learn and “structure” as the main instructional conditions.

It should be noted that these review studies are based on research conducted in western industrialized countries. An overview of school effectiveness studies in developing countries is provided in Scheerens, 2000. An important study carried out in 13 Latin American countries (Willms and Somers, 2001) more or less confirmed some of the central factors from the review studies cited. These authors conclude that, across countries, effective schools were characterized by:

- 1) high level school resources, including a low pupil-teacher ratio, more instructional materials, a library and well-trained teachers;
- 2) classrooms which are not multigrade classes, and where students are not grouped by ability;
- 3) classrooms where children are tested frequently;
- 4) classrooms and schools with a high level of parental involvement; and
- 5) classrooms that have a positive classroom climate, especially with respect to classroom discipline

Obviously in poorer countries there tends to be more variation in basic material and human resources related conditions of schooling, so that these conditions come out more prominent in effectiveness studies.

In three recent “State of the Art” review studies by Reynolds et al. (2013), Muijs et al. (2013) and Hopkins et al. (2013) an overview is given of the most relevant factors in three respective sub-fields: education effectiveness

research (EER), teaching effectiveness research (TE), and school and system improvement (SSI). A summary is provided in Table 3.

EER	TE	SSI
Effective Leadership Academic focus A positive orderly climate High expectations Monitoring progress Parental involvement Effective teaching (time) Staff professional development Pupil involvement	Opportunity to learn Time Classroom management Structuring and scaffolding, including feedback Productive classroom climate Clarity of presentation  Enhancing self regulated learning Teaching meta-cognitive strategies Teaching modelling  More sophisticated diagnosis Importance of prior knowledge	Dimensions of organizational health School based review School development planning Comprehensive School Reform Facets of educational leadership (transformational, instructional, distributed)  Effective systemic reform; see page 15 Hopkins et al., among others, student achievement and teaching quality emphasis.

*Table 3: Effectiveness enhancing conditions referred to in the review studies by Reynolds et al. (2013), Muijs et al. (2013) and Hopkins et al. (2013)*

Once again there is a fair consistency in the factors that are mentioned in the three more contemporary reviews, for examples with respect to core factors like: academic emphasis, time and opportunity, structuring and scaffolding, leadership and monitoring. Moreover, most of these factors also appear in the earlier reviews. In the more recent reviews there is more differentiation and emphasis on classroom level instructional variables, both from the tradition of structured teaching and direct instruction and

from more constructivist orientations (importance of prior knowledge, self regulated learning and teaching meta-cognitive strategies). From this consistency among review studies it might be concluded that school and educational effectiveness research have an established knowledge base. However, two notes of dissonance are to be considered. Firstly, behind this consensus on general characteristics hides considerable divergence in the actual operationalization of each of the conditions. Evidently concepts like “productive, achievement-oriented climate” and “educational leadership” are complex concepts and individual studies tend to vary in the focus that different elements receive.

Scheerens and Bosker (1997, ch. 4) provide an analysis of the meaning of the factors that are considered to work in schooling, as apparent from the questionnaires and scales as used in the actual empirical school effectiveness studies. This work has been taken to a further level of detail by Scheerens et al., (2007). The results of these analyses of variables and instruments, used in research, are provided in Annex 4.

Unlike the agreement on the most important variables in school effectiveness research, reviews of the effect sizes, in the sense of the estimate of the association between a specific factor and educational achievement, show far less consensus. This state of affairs will be elaborated in subsequent sections.

### **Quantitative effects; less consensus about the size of effects**

Meta-analyses compute average effect sizes across individual research studies addressing the association of a certain independent variable and educational achievement. Various coefficients may be used for the estimates. The standardized mean difference (between a treatment and a control group), coefficient  $d$ , and certain correlation coefficients ( $r$ ), are the most common.

Hattie, (2009) provides massive quantitative evidence on the association of numerous school, teacher and teaching variables with student achievement. Average effect sizes for school, curriculum, teacher and teaching factors in terms of the  $d$  coefficient (standardized difference between means) reported by Hattie are .23, .45, .49 and .43 respectively (ibid, pages, 74, 109, 130, 162 and 201). According to Cohen, 1977, effect sizes of .2 are considered small, .5 medium and .8 large. When applying these standards the average effect sizes should be considered as slightly below medium. Still, meta-analyses that are carried out by European authors show effect sizes that are even lower; see for example Witziers, et al. 2003, Scheerens et al., 2007, Seidel and Shavelson, 2007, Creemers and Kyriakides, 2008. By way of illustration some of the results on key variables listed in the three state of the art papers, educational leadership, evaluation and monitoring, learning time, structured teaching and quantity of teaching are compared (further details on how these results were obtained are provided in Annex 1)

School level variables			
	Schee-rens et al., 2007	Hattie, 2009	Creemers & Kyriakides, 2008
Consensus & Cohesion	.02	-	.16
Orderly climate	.13	<b>.34</b>	.12
Monitoring & evaluation	.06	<b>.64</b>	.18
Curriculum/OTL	.15	-	.15
Homework	.07	<b>.30</b>	-
Effective Learning Time	.15	<b>.34</b>	-
Parental involvement	.09	<b>.50</b>	-
Achievement orientation	.14	-	-
Educational leadership	.05	<b>.36</b>	.07
Differentiation	.02	<b>.18</b>	-
Teaching level variables			
	Schee-rens et al., 2007	Hattie, 2009	Seidel & Shavelson, 2007
Time and OTL	.08	<b>.34</b>	.03
Classroom management	.10	<b>.52</b>	.00
Structured teaching	.09	<b>.60</b>	.02
Teaching learning strategies	.22	<b>.70</b>	.22
Feedback & monitoring	.07	<b>.66</b>	.01

Table 4: Results from recent meta-analyses (coefficients are based on the Fisher Z transformation of correlations; as Hattie presents effect sizes in terms of *d*, these are indicated in bold).

According to established scientific standards the effect sizes for the key school and teaching variables are medium when one considers the results by Hattie and small when one considers the other meta-analyses. One of the explanations Hattie (2009, p202) offers for the differenc-

es in effect sizes between his results and those by Seidel and Shavelson is that these latter authors have used only studies that controlled for student prerequisites. This could be seen as a sign that the more Europe based studies used stricter quality controls in selecting studies, and might therefore have more credible results. The other explanation might be that effect sizes in the USA, Great Britain and Australia are higher, perhaps due to greater variability in processes and outcomes.

Some recent meta-analyses, carried out by the author and his associates (Scheerens, 2012, 2013, Hendriks, Steen and Scheerens, 2009) show relatively small effect sizes for school leadership, time and evaluation and assessment (see Annex 2).

The results of these recent meta-analyses further qualify the general consensus that is shown on factors at school and classroom level that “work”. As a matter of fact some of these variables appear to have higher effects than others, and this information will be used in drawing practical implications from this literature, in the second part of this paper. It should be noted that the outcomes are fairly robust with regards to age levels and levels of schooling, in the sense that they are consistent for elementary and lower secondary schools (Scheerens et al., 2007)

### School effectiveness in international comparative studies

In IEA studies and PISA, school, classroom and student level background variables from context questionnaires provide measures that can be associated with student performance. In most studies the school and student level context variables show a fair match with those addressed in school effectiveness research. This is of course a deliberate strategy, as one of the purposes of the international studies is to provide policy relevant explanations on performance differences between schools and countries, which is very similar to the “what works” mission of school



effectiveness research. As an overarching re-analysis and overall review on “what works across countries”, based on these international assessment studies has not been carried out, to my knowledge, some miscellaneous study results are briefly reviewed, before some tentative general trends will be formulated. This material is presented in Annex 3.

Generally, the effect sizes of the school and classroom variables in international comparative assessment studies are even lower than would be expected on the basis of the results from meta-analyses. Annex 3 discusses some of the methodological limitations of these studies, which provide some explanation for these studies having difficulty in detecting school effects, which, even in research studies, show up as relatively small. In this way the results from international studies can be seen as a conservative test of “what works in schooling”. Variables that appear to do the best in surviving this conservative test are: opportunity to learn (match between content covered and content that is tested), disciplinary climate, and use of evaluation and assessment for formative application as well as accountability purposes.

### **Robustness versus “contextual dependency” of the school effectiveness factors**

When considering the school factors that were listed in the above (achievement orientation, cooperation, educational leadership, frequent monitoring, a safe stimulating climate and opportunity to learn), the research literature indicates that such factors are supported both at the elementary school level, as at lower secondary (high school) level. Scheerens et al. (2007) analysed the robustness of these school effectiveness factors, with respect to nationality, age level (primary or secondary education), subject matter area and several methodological study characteristics. They found “a relatively consistent slightly higher effect size for studies carried out in primary, as

compared to studies conducted in secondary schools”, whereas the results appeared to be “less clear-cut for the moderator variables subject matter area and country”. The general picture of their analyses indicated that the effect of the school variables in question were fairly robust, when taking into considerations these context characteristics. Results from PISA give the impression that the factors concerned tend to have a slightly higher impact when science or mathematic achievement is the effect variable, than for reading literacy. A general explanation for these results is that reading literacy effects are less exclusively dependent on within school learning in specific classes, but are also dependent on reading outside lesson hours.

Another impression from the review and meta-analysis literature that one gets is that “good” schooling is particularly beneficial for students from less advantaged socio economic background. This has been noted, for example in studies about the effects of instruction time (Sharp et al. 2007), and school size (Leithwood and Jantzi, 2009).

The effect of school level effectiveness enhancing conditions also depends on the homogeneity of national school systems. Scandinavian countries, particularly Sweden and Norway have relatively small between school variance. In such countries there may be more variance at classroom level. Characteristics of national cultures have also been considered for their impact on educational achievement and effectiveness enhancing conditions. The traditional value of education in Asian cultures has been associated with high student motivation, and greater tolerance for large class size. Hallinger and Kamara (2001) present an interesting case study of the way school leadership and school improvement in Thailand is getting shape, seen against the national cultural background, as defined on the basis of Hofstede’s framework for analyzing cultures.<sup>3</sup>

In a subsequent section reference will be made to structural conditions of national school systems, such as school

<sup>3</sup> Hofstede, (1980) distinguishes four dimensions on which national cultures differ: power distance, uncertainty avoidance, individualism-collectivism and masculinity- femininity.

autonomy, nationally established accountability mechanisms, and the degree of differentiation of the secondary school system.

When discussing the knowledge base on educational and school effectiveness, we should bear in mind that it depends for a very high percentage on studies from western industrialized countries, and, among these, particularly Anglo-Saxon countries. Rare studies in Latin America (Willms and Somers, 2000), Africa, (Fuller and Clarke, 1994) and Asia, (Van der Werf et al., 2001), provide some support for the meaning of the effective school model in such contexts, although cultural contingency is underlined at the same time. Yet, analyses of large scale international data-bases, such as those of PISA and TIMSS, (see the previous section and Annex 3) offer very little support for the universal effectiveness of the school factors considered.

## II. Policy

International comparative assessment studies have created a global competition in educational achievement in core subject matter areas like reading, arithmetic/mathematics and science. This means that in many countries there is increased attention for boosting educational achievement by means of special policy programs, system level levers of educational improvement and financial investments in education. Economists have pointed at impressive economic benefits of better student performance, even when established at the level of the first grades of secondary schools, and alternately at the high costs of lagging behind (cf the OECD report “The high costs of low performance”, OECD, 2010). In this section the system level policy context of school effectiveness will be discussed by presenting a conceptual multi level model, by reviewing the evidence on the effectiveness of system level policy factors and structural characteristics of educational systems, and by considering the levers for reform that are mentioned in a recent OECD report on national systems that have been successful in educational reform. Next,

strategies for school level reform and school improvement will be reviewed.

The conceptual structure of educational effectiveness as a hierarchical system

In Figure 3, (source: Scheerens, 2007) education is depicted as a hierarchical system. In the figure, influence across levels is indicated by the dotted arrows that run from higher levels to lower levels. Such across-level relationships can be interpreted in terms of control, facilitation and buffering from a higher level directed at the core process at the next lower level. Depicting education in this way and qualifying the overall system as hierarchical and loosely coupled has the following implications:

- lower level core processes are seen as being contextualized and controlled by higher levels (the vertical aspect);
- despite this notion of higher level control, lower levels are seen as having considerable discretion over their core processes, in other words considerable autonomy. This is the idea of loose coupling between hierarchical levels, sometimes expressed in more prescriptive terms, like “subsidiarity”; a maxim which states that lower level autonomy should be maximized up to the point beyond which it would become counterproductive. Put differently: This approach would imply that what can be reasonably accomplished at a lower level should not be carried out by a higher level.

The degree of higher level control versus lower level autonomy is an issue of central importance at all levels. At system level it is about effective patterns of functional decentralization, which means that, perhaps dependent on the larger context, certain patterns of centralization in some functional domain (e.g. the curriculum) and decentralization in another domain (e.g. financial management) work best. At school level it is about the degree of participative decision making, or “distributed leadership”, and at



classroom level it refers to the balance between strongly structured didactic approaches and more open teaching and learning situations that are expected to invite self-regulated learning. Structure versus independence is a red line that dominates policy and research agendas in education. A second key element in the representation in Figure 3 is the identification of ecological conditions as a separate class of conditions influencing educational performance. This is done by giving a more explicit place to partially controllable composition effects, and their interaction with more directly malleable variables, such as the school climate. The recognition of this kind of contextual conditions emphasizes the partiality of direct control in education, and in this way underlines the loose coupling between the hierarchical levels, but at the same time focuses the attention on a qualitatively different strand of control measures, namely those of selection, admission, grouping and matching of teachers and sub-groups of students as well as on cultural aspects associated with student and teacher body composition.

Figure 3: (source: Scheerens, 2007) Integrated multi-level model of education; the dotted arrows from one system level to the next represent across level influences; feedback-loops are assumed to run from outcomes at each level to the box containing ecological conditions and active policies at each object level and from lower to higher levels, but these are not shown, to avoid a too complex pattern of arrows (for a detailed description, see text).

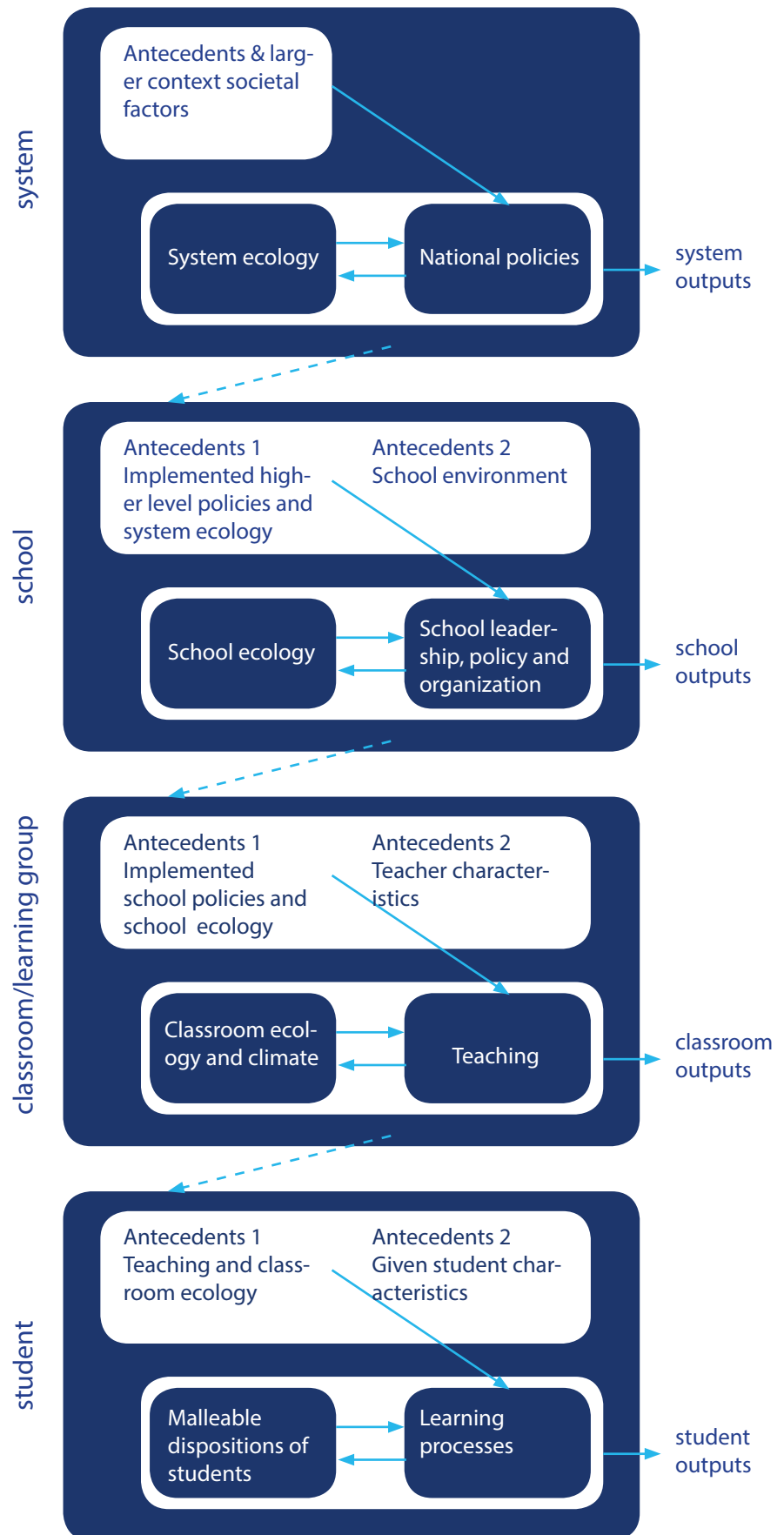


Figure 4: from Scheerens (2007), illustrates how this empty framework can be used as a basis for categorizing variables that have been addressed in empirical research, in this case, school effectiveness research.

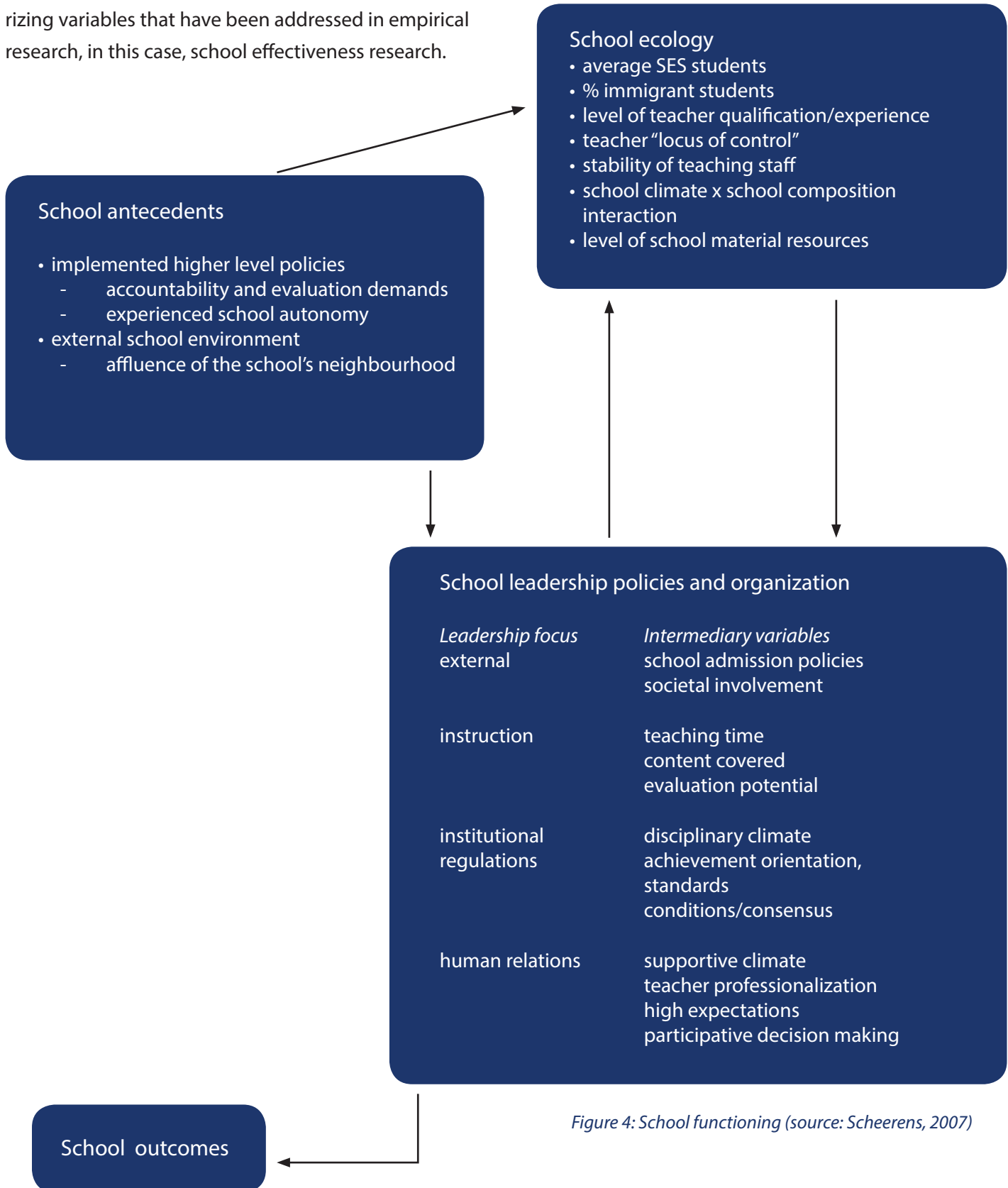


Figure 4: School functioning (source: Scheerens, 2007)

## Empirical studies on the effect of malleable system level variables

A handful of empirical studies have specifically addressed the effect of malleable system level variables on educational achievement. The most important references are: Luyten et al. (2005), Woessmann et al. (2009), Causa and Chapuis (2009), Brunello and Checchi (2006), Scheerens and Maslowski (2008), Jakubowski (2009). Scheerens et al. (2011) provide the following overview of results:

System level variables	
Accountability and a well-developed examination system	Cf. Woessmann et al., 2009; Scheerens et al., 2011 offer an overview. Mostly positive effects of accountability; discussion about side effects of accountability.
School autonomy	Mixed results of school autonomy, mostly not significant (Scheerens & Maslowski, 2008)
Public versus private schools	No effect of private/public, when school composition is taken into account (Luyten et al., 2005)
Stratification (tracked versus comprehensive school systems in secondary education)	Mostly negative effect of separate tracks; see Brunello and Checchi, 2006, Jakubowski, 2009

*Table 5: Illustrative results on accountability, autonomy, choice and stratification as most addressed factors in system level effect studies, cited from Scheerens et al. (2011)*

The most robust of the system level effects are the negative effects of highly differentiated structures of secondary schools, as compared to more comprehensive systems. Key facets of highly differentiated structured are: a relatively low age of first selection in a particular secondary school track (11 or 12 years of age); the number of different secondary school types or categories, a special track for lower vocational education, and high class repetition. It is interesting to note that stratification operates mostly via school and track composition, thus affecting ecology

rather than specific control measures at school level. The variable that is mostly used as a measure of accountability in international studies is the presence of a standard based examination at the end of secondary school (cf Bishop, 1997, Woessmann et al, 2009). However, some studies find that the effect of this variable disappears when the socio economic status of the students is taken into consideration (OECD, 2007, Scheerens et al., 2013). In some cases school accountability policies, for instance in the sense of schools being required to post student achievement results publicly, have also shown positive effects (OECD, 2007).

## Messages from recent studies on successful educational reforms

Recently results from international assessments, including PISA, have been used to identify high performing and successful reforming school systems. Although these studies are retrospective national case studies rather than quantitative analyses, the results are interesting for reflecting on assumed successful levers for reform. Three reports are particularly relevant:

- Strong Performers and Successful Reformers in Education Lessons from PISA for the United States; (OECD, December 2010).
- How the world's most improved school systems keep getting better. (McKinsey & Company, 2010).
- Capturing the leadership premium; How the world's top school systems are building leadership capacity for the future (McKinsey & Company, 2010).

The list of factors that is associated with successful reform, cited from OECD, 2010 is as follows:

- 1) Developing a commitment to education and conviction that all students can achieve high levels
- 2) Ambitious standards aligned with high-stakes gateways and instructional systems (well aligned testing system)

- 3) Developing more capacity at the point of delivery (high quality teachers and school leaders)
- 4) A work organization in which the teachers can employ their potential: management, accountability and knowledge management (flat organization, away from Tayloristic management, school autonomy)
- 5) Institutionalizing improved instructional practice (diagnostic skills, encyclopaedic repertoire, students enthralled, devoted to the improvement of their craft)
- 6) Aligning incentive structures and engaging stakeholders (high stakes examination systems, in collaboration with stakeholders)
- 7) External, professional and parent oriented accountability
- 8) Investing resources where they can make most of a difference (strong teachers aligned to weak students)
- 9) Balancing local responsibility with a capable centre with authority and capacity to act (state sets clear expectations)
- 10) Importance of work based training (transition from school to work)
- 11) Coherence and alignment across levels, policies and practices, and sustained input (policy implementation)
- 12) An outwards outlook of the system (responsiveness)

Scheerens et al. (2013) compared the results of PISA 2000 and 2009, to relate change in system and school level variables to change in reading literacy performance. By looking for the most important changes in school and system level variables in the countries that showed either the highest progress and the strong decrease in reading literacy performance between 2000 and 2009, they also tried to obtain information on effective levers of improvement. However, they found only confirmation for two variables that are more or less in line with the factors shown above, namely school climate and use of evaluation at school level.

el. A striking overall outcome of their study was the high degree of stability in school and system level characteristics between 2000 and 2009.<sup>4</sup>

### **School improvement strategies**

Reform and improvement efforts at school level should be seen as contextualized by these national policies and structural reform measures. Decentralization and increased school autonomy have implications for the degree to which school improvement is partly determined and steered from above school levels, or purely a matter of bottom up development. Accountability policies will have implications for the achievement orientation of schools, and maybe also stimulate “internal accountability” (Carnoy et al., 2003). Educational leadership is affected by both types of policy levers, as well as by the stimulation of school choice.

### **Performance-based approaches to large-scale reform**

Letihwood, Jantzi, and Mascall (2000) state the following properties of the “performance-based approach”:

1. A centrally determined, unifying vision, and explicit goals for student performance, based on the vision.
2. Curriculum frameworks and related materials for use in accomplishing the goals set for students.
3. Standards for judging the quality or degree of success of all students.
4. Coherent, well integrated policies that reinforce these ambitious standards.
5. Information about the organization’s (especially the students’) performance.
6. A system of finance and governance that devolves to the local school site responsibility for producing improvements in system and student performance.
7. An agent that receives the information on organi-

---

<sup>4</sup> In an earlier section it was noted that school effects within countries tended to be less stable over time. It should be noted that in this section, the stability of average student performance across countries as well as national averages on school characteristics are the issue. National averages may be relatively stable, also when, at school level, within countries school effects vary over time.

zational performance, judges the extent to which standards have been met, and distributes rewards and sanctions, with significant consequences to the organization for its success or failure in meeting specified standards.

Leithwood and his co-authors evaluated the impact of five performance-based reform projects (in Kentucky, California, New Zealand, Victoria (Australia), and Chicago) and concluded that only Chicago had demonstrated significant increases in student achievement. They also found that these achievement gains only occurred during the last three of the ten years the program was analysed. During the first six years of the Chicago program “the system operated in decentralized fashion with little functional contact between schools and the district. In other words too little structure characterized the operation” (Hopkins, 2002). During the latter years of the program “five extra district-level functions were developed” and these might explain why students did better during the last years of the programs that were considered in the analyses:

- policy making increasingly supported decentralization
- there was a focus on local capacity building
- a system of rigorous accountability was introduced
- innovation was stimulated
- external support networks were established

Combined arrangements of functional decentralization and accountability that appear to be successful are characterized by centralization on the curriculum and assessment dimension and increased autonomy in areas like personnel management and resource management at school level. “A micro-economic student-level estimation based on data [TIMSS] from 39 countries revealed that positive effects on student performance stem from centralized examinations and control mechanisms, school autonomy in personnel and process decisions..” (Wößmann, 2000).

The example of the Chicago reform program points the attention at two other dimensions that co-determine success:

- pronounced vertical coordination between higher administrative levels and the school level;
- taking into consideration and stimulating local capacity.

Local capacity building has always been one of the main issues in school improvement. School improvement is being considered as a more school-based approach to educational change and innovation as compared to systemic reform as discussed in this section.

### **School based improvement**

School improvement as a field of academic study is seen as a specific branch of the study on educational change. Seminal contributions to the conceptualisation of school improvement are the work of Matthew Miles (Miles, 1998) as well as that of authors like Fullan and McLaughlin and Skillbeck, published in the “International Handbook of Educational Change” (1998) edited by Hargreaves, Lieberman, Fullan and Hopkins. The following characteristics can be seen as the key principles of this orientation to educational change.

- a) The school is the focus of educational change. This means that schools should be analysed as organisations, seen in their local contexts and harbouring the major agents of change, namely teachers.
- b) A strong emphasis on the process dimension of educational change.
- c) The importance of school based “implementation” in the sense of active adaptation or “co-invention” of externally induced changes.
- d) A human relations approach to educational change influenced by group dynamics and the idea of teacher “empowerment”; capacity building and overcoming

professional isolation of teachers. The “counselling” approach of external change facilitators perhaps also fits in this tradition.

- e) An evolutionary “bottom up” view on educational planning and curriculum development.

Within the scientific community active in this field quite a range of emphases can be discerned. These vary from authors like Mitchell and Sackney (2000), who provide a post-modernist view on school improvement and are strongly opposed to accountability and other “mechanistic” approaches, to authors like Reynolds and Hopkins, who relate school improvement to school effectiveness research in emphasising learning and learning outcomes. Still other contributions (e.g. Leithwood et al., 1999, and Hopkins, 2001) integrate school improvement approaches and conceptualisations of systemic reform.

### Comprehensive School Reform

A major break-through in this field is the work of Slavin, who has proposed a “third” way, in addition to the school improvement approach and systemic reform. (Slavin, 1996, 1998). The characteristics of the school improvement approach as described in the above are summarised by Slavin under the heading of “organisational development models”. “Perhaps the dominant approach to school-by-school reform is models built around well-established principles of organisation development, in which school staffs are engaged in an extended process of formulating a vision, identifying resources (such as external assistance, professional development, and instructional materials) to help the school toward its vision, and often locating “critical friends” to help the school evaluate and continually refine its approaches”. Of this approach Slavin says that it is time consuming and expensive. Moreover, he claims that it is only effective for schools that already have a strong capacity for change. “Such schools are ones in which staff is cohesive, excited about teaching, led by a visionary leader

willing to involve the entire staff in decisions, and broadly aware of research trends and ideas being implemented elsewhere.” (p. 1303). Such schools he describes as “seed” schools. A second category of schools Slavin describes as schools intent on doing a better job, but not perceiving the need or having the capability to develop new curricula. According to his categorisation these are schools with good relations among staff and leadership, a positive orientation toward change, and some degree of stability in the school and its district. Finally, as a third category, he refers to schools “in which even the most heroic attempts at reform are doomed to failure. Trying to implement change in such schools is like trying to build a structure out of sand” (ibid 1303). Accordingly he refers to these schools as “sand” schools.

School improvement of the organisational development kind (as we have seen the predominant perspective on school improvement) is considered only feasible in “seed schools”, which he estimates at 5% of all schools in the USA. Sand schools, also about 5% of all schools, would require fundamental changes before they can support any type of school change. The overall majority of schools, according to Slavin, are the brick-schools and they could most efficiently benefit from what he calls comprehensive reform models. His “Success for All” program is an example. Comprehensive reform models provide schools with specific student materials, teachers’ manuals, focused professional development, and relatively prescribed patterns of staffing, school governance, internal and external assessment, and other features of the school organisation. It should be marked that “Success for All” is one of the few improvement projects that has been thoroughly empirically evaluated and has shown to be successful (Slavin, 1996, Scheerens & Bosker, 1997). Similar successes have been reported by Stringfield and others (1995) presenting the idea of schools as “high reliability organisations”. Borman et al. (2003) report results of meta-analyses of CSR programs, indicating small positive effects (effect sizes in the order



of .09 and .15). This study also provides information of concrete programs in the USA that were successful.

It is interesting to note that Slavin's conception (and also its actual realisation in "Success for All") of Comprehensive School Reform models, seems to have returned full circle to the point where, according to Miles, the school improvement movement started its human relations/implementation approach in the 1950's. Namely the discussion on the applicability of externally developed pre-structured innovation programs and curriculum material. The fact that there is evidence that this approach works is very important, and puts a question mark behind the efficiency of forty years of educational innovation based on the less directive, bottom up, social psychological, organisational development approach to school improvement. The Comprehensive School Reform breakthrough came about in the nineteen nineties, and received a boost by the call for "evidence based" educational policy in the United States (which was also followed up in other countries)<sup>5</sup>. It is beyond the scope of this review to assess the development of this approach internationally. As a more anecdotal comment, I might add that in a country like the Netherlands this breakthrough has not happened and is not likely to do so in the future. Here the traditional organizational development "bottom up" approach to school improvement has persisted, held in the saddle by school autonomy and vested interests of an educational support structure thoroughly acculturated in non directive school counselling.

## Chapter 3: Best practices in making schools more effective

### Preface

In the previous chapters school effectiveness has been de-

fined, and foundational issues of the school effectiveness concept have been referred to. With respect to practical applications of the knowledge base a first key issue is the realization that school effectiveness research is about laying bare malleable conditions of schooling, which can be directly applied in practice to improve schools. A second key issue is the contextualization of school effectiveness in system level policy amenable and ecological conditions and the way instructional effectiveness is embedded in school effectiveness. In this third chapter further steps are taken to assist in applying the school effectiveness knowledge for purposes of school improvement.

First of all an attempt is made at further qualification and prioritization of the school factors that appear to "work"; this is done by reconsidering several meta-analyses, and by looking at several other applications and sources of practical knowledge: school quality factors used in law suits, detection of failing schools, and category frameworks for school inspection, used by Inspectorates of education. Key factors are further defined and (in the annex) operationalized.

Secondly, the interrelationship between school effectiveness and teaching effectiveness, as well as the relationship between system level levers for educational reform and school effectiveness are discussed. These analyses illustrate how schools can become more effective through stimulating instructional effectiveness and selecting good teachers, and how schools may be stimulated to enhance their effectiveness in interaction with external constituencies. To an important extent stimulating school effectiveness can be seen as finding adequate reactions towards external stimuli and internally managing good teaching.

In the third place the interrelatedness of school effectiveness enhancing conditions is considered, resulting in pro-

<sup>5</sup> A study by Gross, Booker and Goldhaber, (2009) presents an evaluation of the effects of CSR funding in the USA. The authors note that this special funding has come to a close in 2007, and their study showed no significant effect of the funding of reading literacy performance. As the owners point out, their study did not check on the implementation of the program and schools, and, in this way is not a direct falsification of the CSR approach as a school improvement model.

posals of a more limited set of composite indicators.

Their dynamic application can be considered as alternative substantive strategies for school improvement.

## I. School factors that “work”

### Content

Best practices for making schools more effective involve content and process aspects. Content directly refers to the empirical knowledge base; process relates to dynamic levers for improvement and improvement strategies.

### Weighing the evidence from meta-analyses

School effectiveness research is mostly field research. From the perspective of applicability, this can be seen as an advantage. Another way to express this would be to say that school effectiveness research will tend to have high ecological validity. In a preceding section, when discussing the demarcation between school effectiveness and school improvement, the improvement potential of the key independent variables in school effectiveness research was already mentioned, and underlined by pointing at the malleable nature of these variables. Referring again to the knowledge base on educational and school effectiveness, the question “what works best in schooling” could be answered by a) considering the set of factors on which a

fair consensus among reviewers exists (see the overviews in Figures 1 and 2), and b, by rank ordering these variables according to the average effect size reported in meta-analyses. Any attempt at this kind of synthesis should be seen as tentative, because of the noted variation in effect sizes across meta-analyses, and the fact that it is not possible to capture a moving target, as new results are continuously added to the knowledge base. Nevertheless an attempt at such a tentative synthesis will be made by putting together main results from Marsano (2003), Scheerens et al. (2007) and Hattie, (2009), see Table 6. The results that Marsano presents depend to a large extent on a meta-analysis by Scheerens and Bosker, 1997. Hattie’s results are based on syntheses of numerous meta-analyses for each variable. In a few cases, there was not a straightforward match with variables that were included in Hattie’s synthesis of meta-analyses, and somewhat specific operationalizations were chosen; these are marked and explained in the legend of the table. The variables mentioned in the overview by Marsano are taken as the starting point, and rank-ordered from high to low in their association with student achievement. In the fourth column of the table the average of the three coefficients for each variable are shown. It appears that the original rank ordering by Marsano is preserved in the averages. The effect sizes are rendered in terms of the d- coefficient.

	Marzano, 2003	Scheerens et al. 2007	Hattie, 2009	Average effect size
Opportunity to learn	.88	.30	.39*	.523
Instruction time	.39	.30	.38	.357
Monitoring	.30	.12	.64	.353
Achievement pressure	.27	.28	.43**	.327
Parental involvement	.26	.18	.50	.313
School climate	.22	.26	.34	.273
School leadership	.10	.10	.36	.187
Cooperation	.06	.04	.18***	.093

Table 6: Rank ordering of school effectiveness variables according to the average effect sizes (d-coefficient) reported in three reviews/meta-analyses. \*) operationalized as “enrichment programmes for gifted children”; \*\*) operationalized as “teacher expectations”; \*\*\*) operationalized as “team teaching”.



Of course the labels of the variables are quite general. In Appendix 4, cited from Scheerens et al. (2007) the range of specifications that is behind these general labels are made explicit. The appendix, in this way, gives more flesh and blood to the broad meaning of the variables mentioned in Table 11. At the same time even the general labels provide a relatively clear idea on what aspects of school functioning should be optimized in order to enhance student performance. Opportunity to learn basically refers to a good match between what is tested or assessed in examinations and the content that is actually taught. Instruction time may be expressed in a more global sense as officially available or allocated learning time or more specifically as “time on task”, or “academic learning time”. Monitoring may include various types of school based evaluations, like school based review, school performance feedback, or school aggregate measures of formative assessment at classroom level. Parental involvement might mean the actual involvement of parents with school matters, or the policies by the school to encourage parents to be involved. Achievement pressure refers to school policies and practices that make use of achievement results and performance records, but also to more climate like and attitudinal facets of fostering high expectations of student performance. School climate generally refers to good interpersonal relations at school, but often more specifically to “disciplinary climate” and the fostering of an ordered and safe learning environment. On school leadership many specific connotations are used. Instructional leadership appears to be the most frequently used and successful interpretation in this literature. Cooperation in general terms, often measures with proxy’s like the number of staff meetings, usually has a relatively weak to negligent association with student performance. Only when

cooperation is explicitly task and result oriented somewhat larger effect sizes are found (cf Lomos et al., 2011). When the rank ordering of these results is further contemplated it appears that curriculum variables (opportunity to learn and learning time) predominate. Monitoring could be seen as part of this curricular “syndrome”, but could also be seen as a broader performance lever, which might include teacher appraisal, and schools being part of accountability schemes. The first four highest ranking factors are all to do with a focus of the primary process of teaching and learning at school. The lowest four factors are organizational measures, or “secondary processes”. In the school improvement literature variables like staff cooperation and school leadership are overrated for their importance, when one considers the quantitative evidence on performance effects. An orderly school climate is more like an organizational condition that is directly supportive of the primary process, in the sense that it is about creating a safe and productive learning atmosphere.

### **Correspondence with school factors considered important in practical applications**

Generally the variables that emerge from empirical school effectiveness research are accepted as making sense to teachers and school heads. In applications like legal claims against malfunctioning schools, detecting failing schools and the evaluation of schools by Inspectorates of education, a similar selection of school characteristics is often made. This is illustrated by the following examples.

Table 7 lists variables that are the object of education rights litigation, in the USA, Welner, (2010)

**Variables in education rights litigations (USA) – (Welner, 2010)**

opportunity to learn
credentialed teachers
safe environments
well-maintained buildings
relatively small class size
challenging and engaging class work
available and current text books
computer and internet acces
same performance standards for all schools (including those in poor communities)
no tracking by ability

*Table 7: School variables in US education rights litigation; source, Welner, (2010)*

The context of these litigations is the Title I legislation, implying that schools have the duty to create opportunities to reach the achievement standards that are part of the accountability programs.

The fact that these variables are used in juridical causes underlines that favourable conditions for schooling are recognized as concrete levers for enhancing the educational chances of students, including disadvantaged students.

Table 8: based on Stringfield, 1998 (p. 209-217) lists factors associated with failing schools.

<b>School level</b>	<ul style="list-style-type: none"> <li>- lack of academic focus</li> <li>- teachers working in isolation</li> <li>- academic periods starting late and ending early</li> <li>- lack of coordination between teachers in use of textbooks</li> <li>- bureaucratic leadership, not curriculum or instruction oriented</li> <li>- head teachers passive in teacher recruitment</li> <li>- lack of teacher assessment</li> <li>- no public rewards for students' academic excellence</li> <li>- difficulties in maintaining funding</li> <li>- underutilization of library</li> </ul>
<b>Classroom level</b>	<ul style="list-style-type: none"> <li>- a leisurely pace</li> <li>- minimal planning</li> <li>- low rates of interactive teaching</li> <li>- parts of mandated material not covered in teaching</li> <li>- lack of any sense of academic push</li> </ul>
<b>Student level</b>	<ul style="list-style-type: none"> <li>- low time on task</li> <li>- low opportunity to learn in academic subjects</li> <li>- classes experienced as "intellectual anarchy" (lack of structure)</li> </ul>

*Table 8: Characteristics of failing schools, (Stringfield, 1998)*

To a considerable degree the factors that are listed by Stringfield are negative statements on variables like academic focus, teacher cooperation, instructional leadership, monitoring, time on task and opportunity to learn, that, in their positive formulations, are considered as effectiveness enhancing conditions. In many countries Inspectorates of

Education monitor the quality of schools. In the example illustrated in Table 14 the observation categories from the Dutch Inspectorate of Education are summarized. Inspectorates in other countries are applying similar check lists of school and teaching factors (De Volder, 2012). Again a fair correspondence with the results from empirical school and educational effectiveness research can be observed.

#### Observation categories Dutch Inspectorate of Education

Opportunity to learn
Learning time
Supportive climate
Challenging climate
Structured teaching
Activating students
Teaching learning strategies
Classroom organization
Adaptive teaching

Table 9: Overview of observation categories used during school visits by the Dutch Inspectorate of education. Source: Inspectorate of Education, 2012

Apart from showing correspondence with research results, particularly concerning curriculum quality and school climate, these practical applications understandably have an eye for the availability of basic resources (like well-maintained buildings, availability of text books, computer facilities and libraries). Such variables usually show up with very small effect sizes in industrialized countries, and make more of a difference in what we used to describe as developing countries (Scheerens, 2000). A final observation with respect to the illustrations provided in this section is that the check-lists contain both school level and classroom level conditions. An integrative perspective on school and instructional effectiveness, contextualized by making reference to regional or national policy contexts,

appears to be the most fruitful in educational reform and school improvement practice.

## II. Effective teaching and effective teachers in effective schools

Although effective teaching is not the focus of this report, it would be artificial not to refer to it, since effective schooling is, too a large extent, providing support at school level for optimizing teaching at classroom and individual student level. A good overview of the most relevant variables in teaching effectiveness is provided in Table 10, cited from Brophy (2001).

#### Variables in effective teaching

opportunity to learn
curricular alignment
supportive classroom climate
achievement expectations
cooperative learning
goal-oriented assessment
coherent content; clear explanations
thoughtful discourse
establishing learning orientations
sufficient opportunities for practice and application
scaffolding student's task engagement
modelling learning and self-regulation strategies

Table 10: Variables in effective teaching, from Brophy, 2001

Good, Wiley and Florence, (2009) refer to three latent teaching factors: structure and classroom management, supportive climate and cognitive activation. The integration of these ideas and Brophy's overview of teaching variables is shown in Table 11, based on Klieme, 2012.

<b>Structure and classroom management</b>
Opportunity to learn
Available time
Degree of student involvement
Curriculum alignment
Visible and coherent planning
Goal oriented assessment
Focus on what is important
<b>Supportive classroom climate</b>
Pro-active and supportive classrooms
Caring communities
Appropriate expectations
Help students to exceed
<b>Cognitive activation</b>
Coherent content
Sufficient depth
Thoughtful discourse
Scaffolding students' ideas and task involvement
Understanding at a higher level
Authentic application of concepts in different contexts

Table 11: Latent and manifest teaching variables, adapted from Klieme, 2012.

While teaching effectiveness focuses on teaching processes, teacher effectiveness tries to identify teacher characteristics, like skills, experiences, dispositions and sometimes even personally traits, associated with teaching quality and student achievement. The following type of teacher characteristics will be briefly reviewed: personality traits, formal qualifications and experience, subject matter expertise and knowledge about teaching and learning and pedagogical content knowledge.

### Personality traits

Throughout the history of teacher and teaching effectiveness research personality characteristics of teachers have been investigated as well, looking at variables like: flexibility/rigidity, extraversion/ introversion, locus of control,

self-efficacy, general and verbal intelligence (cf. Brophy, 1983; Darling-Hammond, 1999).

In the nineteen sixties and seventies the effectiveness of certain personal characteristics of teachers was particularly studied. Medley & Mitzel (1963), Rosenshine & Furst (1973) and Gage (1965) are among those who reviewed the research findings. From these it emerged that there was hardly any consistency found between personal characteristics of the teacher like being warm hearted or inflexible on the one hand, and pupil achievement on the other. In a more recent review Darling-Hammond (1999) concludes that effects of general intelligence are inconsistent and small, but that some studies have convincingly demonstrated a positive impact of verbal ability.

### Formal qualifications and experience

Effects of teacher education, usually expressed in terms of formal qualifications, like having a BA or MA degree, or being certified to teach in a specific field, have traditionally been included in studies into "education production functions". In developed, industrialized countries, factors like formal qualifications do not appear to make much of a difference. In developing countries such variables appear to be more often of significant impact. The explanation for this phenomenon is probably that the variation in formal teacher training in developed countries is usually quite limited, and teachers are more or less uniformly equipped to carry out their job. In developing countries teacher preparation is less uniformly distributed. One could say that in developed countries, the impact of teacher education does not come out strongly from cross sectional and comparative studies, because there is a lack of variability in the variable of interest. In Table 12 which combines results from two meta-analyses by Hanushek (1995, 1997), the larger impact of teacher education in developing countries is illustrated.

Input	Industrialized countries % sign. positive associations	Developing countries % sign. Positive associations
Teacher/pupil ratio	15%	27%
Teacher's education	9%	55%
Teacher's experience	29%	35%
Teacher's salary	20%	30%
Per pupil expenditure	27%	50%

*Table 12: Percentages of Studies With Positive Significant Associations of Resource Input Variables and Achievement for Industrialized as Compared to Developing Countries (Sources: Hanushek, 1995, 1997).*

In a way these results are corroborated by the outcomes of studies in the United States about alternative certification of teachers, i.e. other than official full teacher qualifications, as well as studies that have looked at out of field teaching (teaching a subject for which a teacher holds no official qualification). Wayne and Youngs (2003), when summarizing studies by Goldhaber and Brewer (1997 and 2000) noted that for mathematics, results of fully certified teachers were better than for non formally qualified or alternatively qualified teachers. Similar results were not confirmed for other subjects. In a study using state level data from the USA, Darling-Hammond (1999), used a more fine graded scale of teacher qualification, distinguishing between:

- teachers with full certification and a major in their field;
- teachers with full certification;
- teachers less than fully certified;
- uncertified teachers.

She found substantial positive effects for certified teachers and substantial negative effects for uncertified teachers

(correlations in the order of .71 to -.51).

Results of studies which have investigated the effects of teacher experience are not always showing the expected positive effect. According to Darling-Hammond (1999, p. 9) effects are not always significant, nor linear. Effects of experience are particularly visible when teachers with less than 5 years of experience are included in the study.

### Subject matter knowledge and knowledge about teaching and learning

Breaking up the black box of teacher education, the most frequently addressed analytic variables in explaining why some teachers are more effective than others are subject matter mastery and pedagogical knowledge. In the more recent research literature, an interactive construct, combining the two, namely "pedagogical content knowledge" appears to show promising results.

In her review, Darling-Hammond (1999), refers to studies which have correlated teachers' course taking in subject matter areas and scores on subject matter tests to student achievement. She concludes that the former show positive effects more frequently than the latter. Low variability in test scores is seen as the main reason for low and insignificant associations. Subject matter mastery is seen as a basic requirement that is relatively uniformly addressed in initial teacher training. In this sense the explanation of the results on teachers' subject matter mastery is the same as the one already given with respect to overall teacher education effects. Hawk, Coble and Swanson (1985) found that the relationship between teachers' training in science and student achievement was greater in higher level science courses.

Darling-Hammond (ibid) lists some ten studies that indicate that pedagogical training generally has a stronger effect than subject matter mastery. It should be noted that most of the studies she refers to have looked at subject matter related teaching methods. As suggested by Byrne



(1983), effects of subject matter mastery are likely to interact positively with knowledge on how to teach the subject to various kind of knowledge, meaning that the impact of subject matter mastery is augmented by subject matter related didactic knowledge. Wayne and Youngs (2003), on the other hand, present results that show that pedagogical training in language teaching appeared to lower student achievement.

### **Pedagogical content knowledge**

In his seminal article in the *Education Researcher*, Lee Shulman (1986) criticized a sharp division between subject matter mastery and pedagogical skills of teachers. He introduced the concept of pedagogical content knowledge, briefly described as “subject matter knowledge for teaching”. Pedagogical content knowledge is about the selection of topics, useful forms of presentation, analogies, illustrations, examples, explanations, and demonstrations. Pedagogical content knowledge also includes understanding of what makes the learning of specific topics easy or difficult, which includes knowledge about conceptions and misconceptions that students bring to the lesson. The assumption is that “deep knowledge” about the content and structure of a subject matter area forms the crucial precondition for teachers manifesting pedagogical content knowledge in their teaching. Additional components sometimes included in the concept are knowledge on the appropriate use of teaching materials and media, as well as strategic knowledge on the application of teaching strategies. Studies investigating the effect of pedagogical content knowledge are those by Hill et al. (2008) and Baumert et al. (2005). Reviews are provided by Putnam and Borko, (2000), Gess-Newsome and Lederman (2001) and Gess-Newsome, (2009)

Referring back to the multi level framework that encompasses system, school and instructional level effectiveness, effective schooling can be seen as a matter of recruiting effective teachers, stimulating and facilitating effective

teaching, creating favourable contextual conditions and “ecology”, and acting responsively to policy inputs from higher administrative levels and the national context at large. In the final section of this chapter the attention will be on the dynamic process of school improvement, in which the interplay of conditions at different levels is a central issue.

### **Process; or dynamic application of the educational effectiveness knowledge base for school improvement purposes**

#### *How effective teaching variables are stimulated by school level conditions*

For some of the variables identified in the context of school effectiveness straightforward counterparts appear in studies on instructional effectiveness. Opportunity to learn at school level, expressed in terms of the content that is covered in the school curriculum, or the prescription of textbooks aligned with the subject matter required in tests and examinations, clearly structures the actual teaching and implementation of the identified subject matter areas and content categories at classroom level. Likewise, allocated time, according to the school time table sets the stage for instruction time and time on task at classroom level. Bosker and Scheerens (1996) describe this kind of correspondence as the mirroring of school level and instructional level conditions. For most of the other variables the association between school and classroom level conditions is somewhat looser and more indirect, more like facilitation of the classroom level variables by the school level ones. This is the case for clear disciplinary school rules and an orderly and safe classroom climate, and for making available evaluation instruments at school level, for the monitoring of student progress at classroom level. Leadership can have a buffering function, as well as an overall stimulating role on effective teaching. Buffering might imply that the principal takes care of external contacts and administrative burdens that might otherwise keep

the teachers from concentrating on their teaching job. In the study of school leadership effects, indirect effect models are increasingly being applied, in which the principal has impact on intermediary conditions at school and classroom level, which, in their turn, influence student achievement (cf. Hallinger & Heck, 2010, Day et al., 2009, Scheerens, 2012). Examples of such intermediary variables are cooperation between teachers, subject matter alignment, managing the teaching and learning program, and stimulating professional development of teachers. So far the success of these intermediary effect models of school leadership is rather limited, as effect sizes tend to be small to negligible, and effective intermediary conditions appear to vary a lot between studies (Scheerens, 2012). Studies that did show some interesting results in this area, are those by Louis et al. (2010) and Heck and Moriyama (2011).

The idea of school level conditions facilitating the effectiveness of classroom level conditions has a lot of credibility. At the same time this instrumental perspective is constrained by the loosely coupled nature of schools as organizations (Weick, 1976, 2009). As emphasized in the overall multi level framework depicted in Figure 3, the teaching level is rather autonomous and teachers are relatively independent professionals.

Reasoning from the instrumental alignment of school and teaching level conditions has a technological flavour and, from an organizational theoretic perspective can be seen as an example of the configuration hypothesis of contingency theory (internal alignment of different components of the organization- Mintzberg, 1989)). Alternative managerial strategies might be more focused on teacher recruitment, professional development and human resources management on the one hand, and operating via the direct environment of the school, through parent involvement and active involvement with the local environment.

### III. Contextualization of school effectiveness, embedded in national policies and structures

In actual practice the degree to which schools are result oriented and are concerned with improvement of their effectiveness is likely to be shaped by higher level policies and contextual conditions. The higher level could be the school district level, the state level (in federal systems) or the national level. Areas of active higher level policies implemented to enhance school productivity and effectiveness are mostly in the realm of decentralization and providing more autonomy to schools, of creating school choice and market mechanisms in education, and of making schools accountable for their performance.

When it comes to decentralization, it is helpful to realize that educational systems can be differentially decentralized according to specific dimensions of their functioning. This is sometimes indicated as “functional decentralization” (Bray, 1994). In this way the administrative context of schools can be characterized by mixed patterns of centralization and decentralization. Relevant functional areas that are often distinguished are: the curriculum, human resources management, finance and teaching policy. In the latter area, schools tend to have the largest amount of freedom. Less regulation and more autonomy in the area of human and financial resources are often chosen as relatively uncontested areas when systems want to provide more autonomy to schools. The issue of curricular centralization and decentralization is more contested. Despite an overall interest in decentralization some countries, usually those coming from an already rather decentralized tradition, tend to centralize the curriculum by providing more explicit national standards and/or defining a core curriculum (such tendencies have recently been displayed in countries like the UK, the USA, Sweden and the Netherlands). Effect studies show conflicting evidence about the beneficial consequences of curriculum centralization; Woessmann, 2000, noted positive effects, while OECD

(2010), reports positive effects of curriculum decentralization.

Increased school autonomy is expected to stimulate the innovatory potential of schools and the responsiveness of schools with respect to local stakeholders. Free school choice is theoretically seen as the external booster of schools operating under market conditions, and expected to compete for quality to attract students and the support of parents (Chubb and Moe, 1990). The evidence for the effectiveness of free school choice and competition between schools from international studies is weak or missing (e.g. OECD, 2010). Similarly contested are the results of evaluation studies about the effectiveness of Charter schools in the USA (Brookings Institute, 2010, Miron, 2011). One of the explanations behind these results might be that the basic assumption, namely that parents choose schools for their children on the basis of relevant quality indicators is not being fulfilled, because parents use other choice criteria, or do not understand the information provided in quality reports that are provided by schools or above-school level organizations.

One of the interpretations of autonomy is “consumer oriented” accountability, where schools are expected to publicly post achievement results to inform parents and the local community. In fact, in this application increased school autonomy, accountability and school choice all come together.

More generally the relationship of increased school autonomy and accountability policies is interesting, as they might appear as conflicting tendencies, where freedom provided to schools is immediately constrained by standardized restrictions in the form of accountability requirements, centred on school outcomes. As a matter of fact such seemingly conflicting tendencies can be accommodated by the framework of functional decentralization, where school autonomy could be seen as providing more freedom on input and processes facets of school function-

ing, while accountability might imply stricter output control. This combination of freeing process and controlling outcomes confirms to the maxims of “New Public Management”. The type of accountability, inherent in these policies could be indicated as “administrative accountability”, as schools are held accountable by higher administrative levels. The concept of accountability has two main facets: providing evaluative information publicly, and redress for poor performance (Glass, 1972). Incentive schemes attached to administrative accountability may vary from low to high stakes to schools and teachers. The mildest form of requirement could just be that schools post performance data publically, while high stakes accountability could imply financial implications. Finally, two main technical forms of accountability policies can be distinguished, test based accountability and school inspection. Evaluative results on the effectiveness of accountability policies are again mixed, although positive effects are often reported when educational systems that have standardized based examinations are compared to countries which do not have these examinations (Bishop, 1997, Woessmann et al 2009); and sometimes positive effects are found for school publicly posting their achievement results (OECD, 2007).

Yet, of the three system level policy orientations, providing more autonomy to schools, stimulating free school choice and accountability the latter seems to affect school functioning most. The way schools are affected is sometimes described as negative and sometimes as positive. Strategic behaviour of schools, going as far as cheating, is often documented as a negative side effect of accountability policies (Koretz, 2005). Although some studies (e.g. Hanushek and Raymond, 2005) ventilate the message that such strategic behaviour is more of a marginal nature. A very interesting positive interpretation is provided in the study by Carnoy, Elmore and Siskin (2003), when they report evidence that schools, in a context of severe high stakes external accountability, develop what they indicate as “internal accountability”. Internal accountability is described



as “the creation of active problem solving in schools” and “school wide coherence an agreement on expectations for student performance an instructional practice” (ibid, p. 208). A more global explanation for the positive association between evaluation arrangements and performance might be that these provisions commonly enhance the result and outcome orientation in schools (Scheerens and Bosker, 1999).

Apart from educational policy reforms with respect to increased school autonomy, accountability arrangements and choice, specific structural arrangements of educational systems may make a difference. In an earlier section reference was made to repeated research findings that, all other things being equal, comprehensive secondary school systems tend to do better than categorical, tracked systems. A second structural arrangement at state/national level that matters concerns the level of teacher qualifications, seconded by the public appreciation of the teaching profession. The number one show case of the importance of these arrangements is Finland, where teachers are expected to have a Masters degree, are involved in continuous professional development, and have high public esteem (Sahlberg, 2009).

## IV. Conclusion: strategies for improving school effectiveness

### The substantive focus of school level strategies

The school and educational effectiveness knowledge base provide an instrumental orientation to school improvement, meaning that enhancing identified school factors is expected to lead to better student performance. In very broad terms the variables identified in educational effectiveness have to do with the technology of the curriculum (as intended and implemented) and with facets of the organizational climate. In this way one could say that a first broad orientation to school improvement could be labelled as the technology and climate emphasis.

However, it should be noted that schools can choose alternative orientations. A second strategy might be labelled the teacher recruitment and professional development strategy. According to this strategy most of a school’s energy to improve should be focussed at teacher issues, including human resources management. In the third place schools could capitalize on matching and grouping issues. Matching could be both externally oriented, towards the local community, towards higher administrative levels, other schools and to parents, and internally oriented in grouping of students in classrooms and learning groups and assigning teachers to these groups of students.

As noted in the above, applying the knowledge base of educational effectiveness research is closest to the Technology and Climate orientation. The general factors that have been discussed and rank-ordered in previous sections are all candidates to be stimulated. More minute and detailed descriptions of these variables are available in the literature, e.g. Marsano (2003), Scheerens et al. (2007) and Hattie and Alderman, (2012). On the level of strategy choice a more synthetic description of the key factors is considered helpful. The following alternative emphases within the Technology and Climate orientation are distinguished:

- a) Exposure to educational content. This could be seen as a composite of opportunity to learn and instruction time. It expresses the curricular focus and duration of exposure in school curricula and teaching.
- b) Evaluation, monitoring and feedback provisions. Evaluation and feedback can be seen as driving improvement at school and classroom level. Implied facets are clarity of purpose through standards, examination syllabi etc., verification of what students have learned, identification of strengths and weaknesses in content and skills that are mastered, feeding back and diagnosis of outcome patterns, systematic consideration of remedial strategies and setting concrete goals for improvement at student, classroom and school level, in cooperation with other teachers, school principals and

eventual support staff. This latter characteristic could make evaluation/feedback/systematic corrective action the core of task related professional development and teacher cooperation.

- c) Managing the school climate. This involves diverse facets like creating a safe atmosphere, positive interactions, as well as fostering high expectations and pressure to achieve.
- d) Managing the teaching and learning program. Repeated studies, in which more behaviouristic approaches like “direct teaching” were compared to constructivist approaches and where no significant differences in student achievement were found, have inspired analysts to propose more general underlying constructs. One example is the construct of “cognitive activation” (Klieme, 2012), discussed earlier. Another example is the term “focused teaching” coined by Louis et al. Hattie (2009) proposes “active teaching” as an overall construct. Careful attention to lesson planning, variation in structure and independence in learning assignments and keeping students engaged seem to be the core issues in these constructs.
- e) Meta-control as the overriding leadership approach. Meta-control is a concept from control theory, and literary means “control of controllers”. Applied to school leadership this concept emphasizes the notion that schools are professional organization, with teachers as semi-autonomous professionals. Teachers may be metaphorically seen as the prime “managers” of teaching and learning at school. A school leader as a meta-controller is not a laissez-faire leader, but one who sets clear targets, facilitates, and monitors the primary process of schooling from a distance (Scheerens, 2012).

These five strategic angles to the substantive focus of enhancing school effectiveness can be seen as having certain connections. Exposure and evaluation/feedback have a

common element in educational objectives and learning standards. Alignment of what is taught and what is tested is the key issue of opportunity to learn. High expectations and pressure to achieve, as facets of the school climate, likewise need a substantive focus in the form of objectives, standards, assessment instruments and feedback. The educational content dimension, perhaps indicated as the implemented school curriculum, is a core dimension of the teaching and learning programme, next to the ideas on transmission that are more central in concepts like cognitive activation. Managing all of these strategies, as well as their connections, is the task of school leadership as meta-control. Integration of these angles to school improvement, inspired by the educational knowledge base is very close to the approach of Comprehensive School Reform, e.g. Borman et al. (2003).

### **How system level policies could foster these school level strategies**

System level policies and structural characteristics of educational system can be seen as pre-conditions or constraints of school level improvement policies, to which schools need to adapt. More analytically one could ask which system level conditions could be seen as supportive of effective schools and effective school improvement. A third, more “neutral” approach might be to just establish where there are matches between the major system level reform dimensions and structural conditions, as discussed earlier, and the school level improvement strategies.

Following this third approach would favour accountability policies as the best matching system level arrangement for the Technology and Climate orientation to effective school improvement. Accountability policies touch directly on core facets of school functioning, like performance standards, achievement orientation, and perhaps also on the “internal accountability” of schools (see the earlier discussion on the work of Carnoy et al. 2003). As accountability policies are almost inevitably associated with a degree of

centrality in the curriculum, this would emphasize the connection with content exposure and opportunity to learn at school level.

Other system level policies and structural arrangements are more closely associated with alternative orientations. Enhanced school autonomy, as well as strong teacher policies appeal more to teacher recruitment and professional development. Choice and market mechanisms, as well as tracked versus comprehensive school systems, are more associated with admittance, selection and grouping processes at school level.

### **A final note on the process of school improvement**

Among the classic change strategies proposed by Bennis, Benne and Chin (1969), considering school improvement on the basis of the educational effectiveness knowledge base is a clear example of the family of rational empirical strategies. Rational empirical strategies assume a neutral position between “top down” and “bottom up” processes, and innovation strategies that embody these extremes: power coercive strategies on the one hand, and normative re-educative strategies on the other. In actual practice the dominant approach in school improvement, starting from the social agogic approach of Matthew Miles, has been bottom up development. See the section on school improvement in Chapter 2. More recent developments like Comprehensive School Reform programmes and calls for “evidence based” education policy oppose this dominant trend, at least to the degree that room for more external input to school improvement is implied. To this I would like to comment that external input in the form of assessment instruments, guidelines to interpret test scores as well as aligned syllabi and textbooks, are not coerced upon schools but provided as inputs that always allow interpretation and adaptation by professionally autonomous teachers. Writing from a country (the Netherlands) that, according to the OECD indicators, is world champion in school autonomy, and where almost a billion EURO

is annually spent on schools and groups of schools reinventing the wheel for all kinds of complicated educational problems, helped by non directive process counsellors, I am somewhat biased towards more external input to school improvement.

Yet, the critical question of this review should be, do we know enough to provide a strong evidence based input to the practice of school improvement? What we do have is a fair international consensus on the major factors that work in schooling. At the same time there is less consensus on the effect sizes of the major variables. My personal assessment of the effect sizes of the main variables in the realm of what has been called the Technology and Climate orientation is that they are small, when compared to general standards (Cohen, 1988) and medium when they are compared to more arbitrarily standards of “educational significance”. Readers who would like to hear about big effect sizes are referred to Hattie, 2009. I would think that the effect sizes reached with comprehensive school reform programmes, in the order of a d-coefficient of .30, are about what we can obtain by optimizing Technology and Climate at school level. Perhaps some success stories of national reform and improvement as evident from PISA are slightly more optimistic; like the progress made by Germany after the “PISA shock” in 2000, and the improvement of Polish results after integrating vocational and general secondary school tracks.

In a way the alternative orientations, the Teacher Recruitment and Professional Development perspective and the Matching and Grouping orientation might offer additional and maybe even stronger effects. The evidence for this is more coincidental. If one looks at the excellent performance of Finland, this could be seen as strong evidence for this teacher centred orientation. Although fundamental, improving initial training and professional development of teachers are at best very slow and time consuming levers of educational improvement. The potential of the Matching and Grouping orientation might be inferred from the

strong impact of school composition (in terms of the school average socio economic status of the students home background), as established from, among others, PISA data. Yet, optimizing school composition, through selective student intake policies, would usually be considered as stimulating “excellence” at the cost of equity in schooling.

The way the success of schools and students depends on the socio economic background of the students and the school and classroom composition give reason for a more profound relativisation of the degree of malleability in education. Particularly telling are the degrees to which favourable impact of variables like disciplinary climate depend on school composition, in most countries (Luyten et al., 2005, OECD, 2010).

### **International schools as the context of application**

Authors describing the specific context of international schools, and the role of international agencies governing these schools, such as the IBO (International Baccalaureate Organization), underline the diversity of school types and school contexts (Fertig 2007, Hayden and Thompson, 1995, and Hayden, 2011). The specific nuance that was expressed between support for the robustness of the school effectiveness enhancing factors on the one hand, and diversity between countries and national cultures, on the other, should therefore also apply to the use of this knowledge in international schools.

Two other issues that are dealt with in the cited studies on international schools that come to the fore are the humanist tradition of child centred education in the international schools and the importance of learning from social and national diversity. The confrontation between behaviourist and cognitive (or constructivist) orientations in teaching is discussed in the article by Fertig. On this issue it would seem that the hard edge of this debate has worn off over time, and that the perspectives are increasingly being used together and integrated. In the general principles of effective schooling discussed in this paper aspects of both approaches have a place (see particularly the contributions with respect to effective teaching).

Finally, the evidence base for teaching heterogeneous school populations is rising, given the growing cultural diversity in many countries. In Europe this challenge, among others, has been met with a call for teaching European citizenship. This orientation leads to an elaboration of school effectiveness thinking, by considering a broader range of outcomes (such as citizenship competencies) as well as a broader range of relevant school conditions (such as informal learning at school from the way the school is run as an organization, the way intercultural relationships are managed and school rules and internal democracy are getting shape, Scheerens, et al. 2009)

## References

- Baker D.P., Fabrega, R., Galindo, C and Mishook, J. (2004) Instructional time and national achievement: cross-national evidence. *Prospects*, XXXIV, 311- 334.
- Bangert- Drowns, R.,L., (1991) Effects of frequent classroom testing. *Journal of Educational Research*, 85(2), 89-99.
- Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A. & Morgan, M.T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research* , 61, 213-23.
- Baumert, J., Blum, W., Brunner, M., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Neubrand, M. and Tsai, Y-M. (2005). *Professional Knowledge of Teachers, Cognitively Activating Instruction, and the Development of Mathematical Competence (COACTIV)*. Presentation at the OECD workshop of the Networks A and C on Teaching and Learning, Reykjavik, October, 2005.
- Baumert, J., Lüdtke, O., & Trautwein, U. (2006). Interpreting Effect Sizes in Large-Scale Educational Assessments. Internal paper Max Planck Institute for Human Development, Berlin.
- Bennis, W.G., Benne K.D., & Chin, R. (1969). *The Planning of Change*. London: Holt, Rinehart & Winston.
- Bishop, J. (1997). The effect of National Standards and Curriculum-Based Exams on Achievement. *The American Economic Review*, 87(2), 260-264.
- Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: a meta-analysis. *Review of Educational Research*. 73, 125-230.
- Bosker, R. J. (2011) *From Educational Effectiveness to Evidence Based Education*. Key Note Address, ICSEI, Cyprus.
- Bosker, R. J., (1997) An international comparative school effectiveness study, using literacy data. Chapter 7 in:
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Bosker, R.J. (1990). Theory development in school effectiveness research: In search for stability of effects. In P. van den Eeden, J. Hox & J. Hauer (eds.), *Theory and model in multilevel research: Convergence or divergence?* Amsterdam: SISWO.
- Bosker, R.J., & Scheerens, J. (1999). Openbare prestatiegegevens van scholen; nuttigheid en validiteit. *Pedagogische Studiën*, 76(1), 61-73.
- Bosker, R.J., & Scheerens, J. (1994). Alternative models of school effectiveness put to the test. In R.J. Bosker, B.P.M. Creemers & J. Scheerens (eds.), *Conceptual and Methodological Advances in Educational Effectiveness Research*, pp. 159-180. Special issue of the *International Journal of Educational Research*, 21(2).
- Bosker, R. J., Guldemon, H.G., Hofman, R.,H., & Hofman, W. H. A. (1989). De stabiliteit van schoolkwaliteit. In J. Scheerens en J.C. Verhoeven (Eds.), *Schoolorganisatie, beleid en onderwijskwaliteit*. Lisse: Swets and Zeitlinger.
- Bray, M. (1994). Centralization/decentralization and privatization/ publicization: conceptual issues and the need for more research. In W.K. Cummings & A. Riddell (eds.), *Alternative Policies for the Finance, Control, and Delivery of Basic Education* [pp. 817-824]. Special issue of the *International Journal of Educational Research*, 21(8).
- Brookings Institute (2010): *Charter Schools: A report on rethinking the federal role in education*. Washington.
- Brophy, J. (1983). Conceptualizing student motivation. *Educational Psychologist*, 18(3), 200–215
- Brophy, J. (2001). Generic Aspects of Effective Teaching. In Wang, M.C. and Walberg, H.J. *Tomorrow's Teachers*. McCutchan Publishing Company.
- Brunello, G., & Checchi, D. (2006). *Does School Tracking*



- Affect Equality of Opportunity?* IZA Working Paper no. 2348. Bonn.
- Burns, M., K. and Symington, T., (2002) A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, 40, 437 -447.
- Byrne, C.J. (1983). *Teacher knowledge and teacher effectiveness: A literature review, theoretical analysis and discussion of research strategy*. Paper presented at the meeting of the Northwestern Educational Research Association, Ellenville, NY.
- Carnoy, M., Elmore, R., & Siskin, L., S. (2003). *The new accountability. High schools and high stakes testing*. New York: Routledge.
- Causa, O., & Chapuis, C. (2009). *Equity in Student Achievement Across OESO Countries: An investigation of the role of policies*. OESO Economics Department Working Papers, No. 708. OESO Publishing. Doi: 101787/223056645650
- Chin, J. M-C. (2007). Meta-analysis of transformational school leadership effects on school outcomes in Taiwan and the USA. *Asia Pacific Education Review*, 8(2), 166-177.
- Chubb, J.E., & Moe, T.M. (1990). *Politics, Markets and American Schools*. Washington, DC: Brookings Institute.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cotton, K. (1995). *Effective schooling practices: A research synthesis*. 1995 Update. School Improvement Research Series. Northwest Regional Educational Laboratory.
- Creemers, B.P.M. & Kyriakides, L., (2012) *Improving quality in education. Dynamic approaches to school improvement*. London and New York: Routledge.
- Creemers, B.P.M. (1994). *The Effective Classroom*. London: Cassell.
- Creemers, B.P.M., & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. London and New York: Routledge.
- Darling-Hammond, L. (1999). *Teacher Quality and Student Achievement: A Review of State Policy Evidence*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., Brown, E., Ahtaridou, E., & Kington, A. (2009). *The Impact of School Leadership on Pupil Outcomes*. Nottingham, UK: The National College for School Leadership.
- De Volder, I., (2012). *Externe schoolevaluaties in Europa (External school evaluations in Europe). A comparative study*. Antwerpen, Apeldoorn: Garant.
- Fertig, M. (2007). International school accreditation. Between a rock and a hard place? *Journal of Research on International Education*, 6, 333-348.
- Fraser, B.J., Walberg, H.J., Welch, W.W., & Hattie, J.A. (1987). Syntheses of educational productivity research. Special Issue of *the International Journal of Educational Research*, 11(2).
- Fuchs, Lynn S. & Fuchs, Douglas. (1985). *A Quantitative Synthesis of Effects of Formative Evaluation on Achievement*. [Washington, D.C.] : Distributed by ERIC Clearinghouse, <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED256781>.
- Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules and pedagogy. *Review of Educational Research*, 64, 119-157.
- Gage, N. (1965). Desirable behaviors of teachers. *Urban Education*, 1, 85-95.
- Gess-Newsome, J., and N.G. Lederman. (2001) Examining *Pedagogical Content Knowledge: The Construct and its Implications for Science Education*, *Contemporary Trends*

- and Issues in Science Education. - See more at: [http://www.idra.org/IDRA\\_Newsletter/August\\_2009\\_Actionable\\_Knowledge/Pedagogical\\_Content\\_Knowledge/](http://www.idra.org/IDRA_Newsletter/August_2009_Actionable_Knowledge/Pedagogical_Content_Knowledge/)
- Glass, G.V. (1972). The many faces of educational accountability. *Phi Delta Kappan*, 53, 636-639.
- Goldhaber, D.D., & Brewer, D.J. (1997). Evaluating the effect of teacher degree level on educational performance. In W. J. Fowler (ed.), *Developments in school finance*, 1996, (pp. 197-210). Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Goldhaber, D.D., & Brewer, D.J. (2000). Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129-145.
- Good, Wiley & Florez (2009). Effective Teaching: An Emerging Synthesis. In G. Dworkin (Eds.), *International Handbook of Research on Teachers and Teaching*. New York: Springer, pp. 803-816.
- Gray, J., Jesson, D., Goldstein, H., Hedges, K., & Rasbash, J., (1995). A multi-level analysis of school improvement: Changes in school's performance over time. *School Effectiveness and School Improvement*, 6, 97-114.
- Grisay, A. (1996). *Evolution des acquis cognitifs et socio-affectifs des élèves au cours des années de collège*. Liège: Université de Liège.
- Gross, B, Booker, K. T., and Goldhaber, D. (2009) Boosting student achievement: The effect of Comprehensive School Reform on student achievement. *Educational Evaluation and Policy Analysis* June 2009 vol. 31 no. 2 111-126.
- Hallinger, P., & Heck, R.H. (2010) Leadership for learning: Does collaborative leadership make a difference? *Educational Management Administration & Leadership*, 38(6), 654- 678.
- Hallinger, P., & Kantamara, P. (2001). Exploring the cultural context of school improvement in Thailand. *School Effectiveness and School Improvement*, 12(4), 385-408.
- Hanushek, E.,A., & Raymond, M.E., (2005) *School Accountability and Student Performance*. Washington University.
- Hanushek, E.A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *The World Bank Research Observer*, 10(227-46).
- Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: an update. *Educational Evaluation and Policy Analysis*, 19, 141-164.
- Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (Eds.), (1998) *International Handbook of Educational Change*. Dordrecht, Boston, London: Kluwer.
- Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.
- Hattie, J. and Alderman, E.,M. (2012), Eds. *International guide to student achievement*. London, New York: Routledge
- Hawk, P.P., Coble, C.R. and Swanson, M. (1985). Certification: It does matter. *The journal of teacher education*, 36(3), p. 13-15.
- Hayden, M., (2011) Transnational spaces of education: the growth of the international school sector. *Globalization, Societies and Education*. 9, 211- 224.
- Hayden, M., and Thompson, J. (1995) International schools and international education; a relationship renewed. *Oxford Review of Education* 21, 3, 327-345.
- Heck, R. H. & Moriyama, K.(2010). Examining relationships among elementary schools' contexts, leadership, instructional practices, and added-year outcomes: a regression discontinuity approach. *School Effectiveness and School Improvement*, 21(4), 377-408.
- Hendriks, M., Scheerens, J., and Steen R. Meta-analyses of

- the effects of school and classroom evaluation and assessment on student outcomes. Unpublished manuscript University of Twente, The Netherlands.
- Hill, H., Ball, L., and Shilling, G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, 2008, Vol. 39, No. 4, 372–400.
- Hill, P.W., & Rowe, K.J. (1996). Multilevel modelling in school effectiveness research. *School Effectiveness and School Improvement*, 7(1), 1–34.
- Hofstede, G., (1980) *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.
- Hopkins, D. (2002). Educational Innovation: Generic Lessons Learned from (a) Regional Practice. Keynote Paper Presentation at the DECIDE International Seminar, University of Twente the Netherlands.
- Hopkins, D., Stringfield, S., Harries, A., Stoll, L., and Mackay, T., (2012) School and System Improvement: to appear in: *School Effectiveness and School Improvement*, 24.
- Jakubowski, M. (2009). Linking International Surveys of Student Achievement. Presentation at EU Conference Stockholm, November, 2009.
- Janssens, F.J.G. (2007). Supervising the quality of education. In W. Böttcher & H.G. Kotthoff (Eds.), *Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung* [School inspection: Evaluation, accountability and quality development]. Münster: Waxmann.
- Kim, S.,E., (2005) *Effects of implementing performance assessment on student learning: Meta- analysis using HLM*. Unpublished PhD. The Pennsylvania State University, PA.
- Klieme, E., (2012) Qualities and effects of teaching. Towards a conceptual theory of teaching. Key note address. EARLI SIG meeting, Educational Effectiveness, Zuerich, 23 August.
- Koretz, D. (2005) *Alignment, High Stakes, and the Inflation of Test Scores*. CSE Report 65. Harvard Graduate School of Education.
- Leithwood, K., Jantzi, D., & Mascal, B. (2000). *Large-scale reform: What works*, Unpublished manuscript. Toronto: OISE, University of Toronto.
- Leithwood, K., Jantzi, D., (2009) A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research*. 79, 464 –490.
- Levine, D.K., & Lezotte, L.W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, WI: Nat. Centre for Effective Schools Research and Development.
- Lomos, C., Hofland, R., & Bosker, R. (2011) Professional communities and learning achievement- A meta-analysis. *School Effectiveness and School Improvement* (22) 121–148.
- Louis, K.S., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, 21(3), 315 –336.
- Luyten, H., Scheerens, J., Visscher A/, Maslowski, R., Witziers, B., and Steen, R.. (2005) *School factors related to quality and equity*. Results from PISA 2000. Paris: OECD.
- Luyten, J.W. (1994). *School effects: stability and malleability*. Enschede: University of Twente, dissertation.
- Marzano, R.J. (2003) *What works in schools. Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R.J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.



- Marzano, R.J., Waters, T., & McNulty, B.A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development
- McKinsey & Company, (2010) *How the world's most improved school systems keep getting better*. McKinsey & Company.
- McKinsey & Company, (2010). *Capturing the leadership premium* (How the world's top school systems are building leadership capacity for the future). McKinsey & Company.
- Medley, D. & H. Mitzel (1963). Measuring classroom behavior by systematic observation. In: N.L. Gage (ed.), *Handbook of research on teaching*. Chicago: Rand McNally.
- Miles, M, (1998). Finding Keys to School Change: A 40-Year Odyssey. In A. Hargreaves, A. Liebermann, M. Fullan & D. Hopkins (Eds.), *International Handbook of Educational Change*. Dordrecht, Boston, London: Kluwer.
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs: Prentice Hall.
- Miron, G. (2011). Review of "Charter Schools: A Report on Rethinking the Federal Role in Education." Boulder, CO: National Education Policy Center. Retrieved [date] from <http://nepc.colorado.edu/thinktank/review-charter-federal>.
- Mitchell, C., & Sackney, L. (2000). *Profound improvement: building capacity for a learning community*. Lisse: Swets & Zeitlinger.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & R. Ecob (1988). *School matters: the junior years*. Somerset: Open Books.
- Muijs, D., Creemers, B., Kyriakides, L., Van der Werf, G., Timperley, H., & Earl, L. (2012). Teaching Effectiveness. A State of the Art Review. To appear in: *School Effectiveness and School Improvement*, 24.
- OECD (2007). PISA 2006. Science Competencies for tomorrow's world. Paris: OECD Publishing.
- OECD (2010). *PISA 2009 Results Volume I, What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*. Paris: OECD Publishing.
- OECD, (2010). *Strong Performers and Successful Reformers in Education Lessons from PISA for the United States*; Paris: OECD Publishing.
- OECD (2010). *The high costs of low performance*, Paris: OECD Publishing
- Postlethwaite, T., N., and Ross, K., (1992) *Effective schools in reading. Implications for educational planners*. The Hague: IEA
- Purkey, S.C. & M.S. Smith (1983). Effective schools: a review. *The Elementary School Journal*, 83, 427-452.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning. *Educational Researcher*, 29 (1), 4-1.
- Reynolds et al. (2013) Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., Van Damme, J., Teddlie, C., & Stringfield, S. (2012). Educational Effectiveness Research (EER): A state of the Art Review. to appear in: *School Effectiveness and School Improvement*, 24.
- Richard, F.D., Bond, C.F., Jr., & Stokes-Zoota, J.J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7, 331-363.
- Robinson, V.M.J., Lloyd, C., & Rowe, K.J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership type. *Educational Administration Quarterly*, 44(5), 635-674.
- Rosenshine, B. & N. Furst (1973). The use of direct observations to study teaching. In: R.M. Travers (ed.), *Second*

- Handbook of Research on Teaching*. Chicago: Rand McNally.
- Sahlberg, P., (2011) *Finnish lessons*. New York: Teachers College Press.
- Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London: OFSTED.
- Scheerens, J. & B.P.M. Creemers (1989). Conceptualizing school effectiveness. In: *Development in school effectiveness research*. Special issue of the *International Journal of Educational Research*, 13(7). Oxford: Pergamon Press.
- Scheerens, J. (1990). School Effectiveness and the Development of Process Indicators of School Functioning. *School Effectiveness and School Improvement*, 61-80. Lisse: Swets & Zeitlinger.
- Scheerens, J. (1992). *Effective Schooling, Research, Theory and Practice*. London: Cassell.
- Scheerens, J. (2000). *Improving school effectiveness*. Paris: UNESCO, IIEP, Fundamentals of Educational Planning series no. 68.
- Scheerens, J. (2007) Conceptual framework for the PISA 2009 background questionnaires. Internal report, University of Twente.
- Scheerens, J. (2012a) (ed.) *School Leadership Effects Revisited. Review and Meta-Analysis of Empirical Studies*. Dordrecht, Heidelberg, London, New York: Springer
- Scheerens, J. (2013) (ed.) *Productive Time in Education. A review of the effectiveness of teaching time at school, homework and extended time outside school hours*. Enschede: University of Twente.
- Scheerens, J., & Bosker, R.J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Elsevier Science Ltd.
- Scheerens, J., & Maslowski, R. (2008). Autonomie des établissements scolaires: des moyens à la recherche d'un objectif? *Revue française de pédagogie*, 164, 27-36.
- Scheerens, J., (Ed.) (2012). *School leadership effects revisited. Review and meta-analysis of empirical studies*. Dordrecht, Heidelberg, London, New York: Springer
- Scheerens, J., Glas, C., & Thomas, S., M., (2003) *Educational evaluation, assessment and monitoring. A systemic approach*. Lisse: Swets & Zeitlinger.
- Scheerens, J., Glas, C.A.W., Luyten, H., & Jehangir, K. (2013). *System Level Correlates of Educational Achievement*. Manuscript, University of Twente.
- Scheerens, J., Luyten, H., & Van Ravens, J. (2011) *Perspectives on Educational Quality. Illustrative Outcomes on Primary and Secondary Education in the Netherlands*. Research Briefs, Springer.
- Scheerens, J., Luyten, H., Steen, R., & Y. Luyten-de Thouars (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: University of Twente, Department of Educational Organisation and Management.
- Scheerens, Jaap (Ed.) (2009). *Informal Learning of Active Citizenship at School* (Lifelong Learning Book Series, 14). New York: Springer (
- Seidel and Shavelson, 2007 Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research* 77(4), 454-499.
- Sharp, C., Chamberlain, T., Morrison, J., Filmer-Shankey, C.(2007). *Playing for Success. An evaluation of its long term effects*. London: National Foundation for Educational Research
- Shulman, L.S. (1986). *Those who understand: Knowledge*

- growth in teaching. *Educational Researcher*, 15, 4–14.
- Slavin, R.E. (1996). *Education for all*. Lisse: Swets & Zeitlinger.
- Slavin, R.E. (1998). Sands, Bricks, and Seeds: School Change Strategies and Readiness for Reform. Hargreaves, A., Lieberman, A., Fullan, M., And Hopkins, D. *International Handbook of Educational Change*.: Dordrecht, Boston, London: Kluwer.
- Stringfield, S. (1998). An anatomy of ineffectiveness. In: L. Stoll & K. Myers (Eds.), *No quick fixes: Perspectives on schools in difficulties*. (pp. 209-221). London: Falmer.
- Stringfield, S.C., & Slavin, R.E. (1992). A hierarchical longitudinal model for elementary school effects. In B.P.M. Creemers & G.J. Reezigt (Eds.), *Evaluation of Educational Effectiveness* [pp. 35-39]. Groningen: ICO.
- Thomas, S.,M., Peng, W.,J., & Gray, J., (2010) Modeling patterns of improvement over time: value added trends in English secondary schools. *Oxford Review of Education*. 33, 261-295.
- Tymms, Merrell, and Henderson (1997) Tymms, P., Merrell, C., & Henderson, B. (1997). The first year at school. A quantitative investigation of the attainment and progress of pupils. *Educational Research and Evaluation*, 3(2), 101-118.
- Vermeer, N., en Van der Steeg, M., (2011) Onderwijsprestaties Nederland in Internationaal Perspectief. CPB Achtergronddocument bij CPB Policy Brief 05, 2011. Den Haag: CPB
- Wayne, A. J. and Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89-122.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Weick, K.E. (2009). *Making sense of the organization*. Volume Two. Chichester: Wiley.
- Welner, K. G. (2010). Education Rights and Classroom Based Litigation: Shifting the Boundaries of Evidence. In Allan Luke, Greg Kelly, and Judith Green (eds.), *Review of Research in Education* 34, Washington, DC: American Educational Research Association.
- Werf, G. van der, & Guldemon, H. (1995). *Omvang, stabiliteit en consistentie van schooleffecten in het basisonderwijs*. Paper presented at the annual conference of the Dutch Association for Educational Research (Onderwijs-researchdagen), Groningen.
- Werf, G. van der, Creemers, B., & Guldemon, H., (2001). Parental involvement in primary education in Indonesia. *School Effectiveness and School Improvement*, 12, 4, 447-466.
- Willms, J.D., & Somers, M.-A. (2000). *Schooling outcomes in Latin America*. Report prepared for UNESCO-OREALC and the Laboratorio Latinoamericano de la Calidad de la Educación [The Latin American Laboratory for the Quality of Education].
- Witziers, B., Bosker, R.J., & Krüger (2003). Educational leadership and student achievement: the elusive search for an association. *Educational Administration Quarterly*, 39(3), 398-425.
- Woessmann, et al., 2009 Woessmann, L., Luedemann, E., Schuetz, G. & West, M.R. (2009). *School Accountability, Autonomy and Choice around the World*. Cheltenham, UK/ Northampton, MA, USA: Edward Elgar.
- Wößmann, L. (2000). *Schooling resources, educational institutions, and student performance: The international evidence*. (Kiel Working paper No. 983). Kiel, Germany: Kiel Institute of World Economics.

## ANNEX 1: How the results for table 4 were computed

For educational leadership Hattie (ibid p. 82) reports a coefficient  $d$  of .36, which, expressed as a correlation coefficient is .18; Scheerens et al. 2007, and Creemers and Kyriakides, 2008, p. 202 report effects (correlations) of .05 and .07 respectively. For feedback and monitoring at classroom level from Hattie's results an average effect of .33 (correlation) can be computed. Scheerens et al., 2007 and Seidel and Shavelson found effects of .07 and .01 respectively. Using structured teaching as an overarching label, Hattie reports an effect size comparable to a correlation of .30 for direct teaching. Creemers and Kyriakides found an effect size of .17 for "quality of teaching", while Scheerens et al and Seidel and Shavelson (ibid) report effects of .09 and .02 respectively. Variables associated with quantity of teaching, time on task and opportunity to learn show effect sizes of .17 (Hattie, p. 184, Creemers and Kyriakides, p.201), but .08 and .03 in the studies by Scheerens et al 2007 and Seidel and Shavelson, 2007. A table showing these and more comparisons of results from meta-analyses is provided in Scheerens et al., 2011, p.144, this table is reproduced as table 4.

## ANNEX 2: Further results of meta-analyses

As a further illustration of the divergence in outcomes between meta-analyses, three overviews of effect sizes, concerning instruction time, school leadership and evaluation/feedback, are presented in the tables below.

Meta-analysis by	Time described as	Mean Effect size (Cohen's d)	Mean Effect size (Correlation coefficient r)	Number of studies	Number of replications
Fraser et al., 1987 (1)	Instructional time	d =.36	r=.18		
Fraser et al., 1987 (2)	Engaged time	d =.83	r =.38	7827	22155
Fraser et al., 1987 (2)	Time on task	d =.88	r =.40		
Scheerens et al., 2007	Learning time	d =.31	r =.15	30	111
Creemers & Kyriakides, 2008	Quantity of teaching	d =.33	r =.16	18	
Hattie, 2009	Time on task	d =.38	r =.19	100	136
Hattie, 2009	Decreasing disruptive behavior	d =.34	r =.17	165	416
Marzano, 2000	Classroom management	d =. 52	r = .25	100	

Table 2.1 Overview of effect sizes of "time"

Meta-analysis by	Leadership concept	Mean Effect size (correlation)
Scheerens et al., 1997	School leadership	r = .04
Witziers, Bosker and Krüger, 2003	School leadership	r = .02
Marzano, Waters and McNulty, 2005	Generalized school leadership	r = .25
Chin, 2007	Transformational leadership	r = .49
Robinson, Lloyd and Rowe, 2008 (1)	Instructional leadership	r = .21
Robinson, Lloyd and Rowe, 2008 (2)	Transformational leadership	r = .06
Creemers and Kyriakides, 2008	School leadership	r = .07
Hattie, 2009	School leadership	r = .18
Scheerens et al., 2007	School leadership	r = .06
Scheerens, 2012	School leadership (indirect effect models)	r = .06

Table 2.2: Summary of results from meta-analyses on school leadership; effect sizes are rendered as correlations between school leadership and student achievement.

Review studies and meta-analyses concerning evaluation, assessment en feedback are, among others, those by Black & Wiliam (1998), Hattie & Timperly (2007), Shute (2008), Hattie (2009) and Wiliam (2011). In the tables below, result from the following meta-anlsyes have been summarized: Fuchs & Fuchs, 1985, Kim, 2005, Hattie, 2009, Burns & Symington, 2002, Bangert et al., 1991 en Kluger and DeNisi, 1996.

Formative evaluation		
Variable description	Effect size (d)	Reference
Formative evaluation	.70	Fuchs & Fuchs, 1985
Formative evaluation	.39	Kim, 2005
Formative evaluation of programs	.90	Hattie, 2009
Formative evaluation followed by an intervention	1.10	Burns & Symington, 2002

Table 2.3: Results of meta-analyses on formative evaluation

Assessment and testing		
Variable description	Effect size (d)	Reference
Prestatie toetsing	.39	Kim, 2005
Frequent toetsen	.23	Bangert- Drowns, 1991

Table 2.4: Results of meta-analyses on assessment and testing

Feedback		
Variable description	Effect size (d)	Reference
Feedback with cues	1.10	Hattie, 2009
Feedback with reinforcement	.94	Hattie, 2009
Feedback, computer assisted	.52	Hattie, 2009
Feedback, overall	.26	Bangert- Drowns et al., 1991
Feedback, overall	.38	Kluger & DeNisi, 1996
Feedback + correct response	.43	Kluger & DeNisi, 1996
Feedback on changes earlier tasks	.55	Kluger & DeNisi, 1996
Feedback and goal setting	.51	Kluger & DeNisi, 1996

Table 2.5: Results of meta-analyses on feedback

Again these results show sizeable differences between the various meta-analyses, for each concept. The results on evaluation, testing and feedback, presented in the above are considerably higher than for time and leadership. This is probably due to the fact that most of the results on the evaluation, assessment and feedback variables were computed in micro level studies at the classroom level.

But how should the size of the effects be interpreted in an educational context? According to Cohen's standards for interpreting effect sizes, most of the effect sizes are small to medium. It should be noted however, that several authors argue that Cohen's standards are to be considered as too conservative, and do not match the practical significance of malleable school variables. Richard, Bond, and Stokes-Zoota (2003; cited by Baumert et al., 2006) found a mean correlation of  $r = .21$  in their meta-analysis of meta-analyses in social psychology, and proposed a modification of Cohen's classification, considering a correlation of .30 to indicate a large effect (p. 339). Baumert, Luedtke and Trautwein (2006) propose the learning gain during one school year as a realistic standard to express effects of schooling. They cite several studies that indicate that this learning gain has the magnitude of about  $d = .30$ . These authors also discuss a method to compute effect sizes developed by Tymms, Merrell, and Henderson (1997), which, when applied to a practical example, suggests that effect sizes of about  $r = .15$  to  $.20$  (small, according to Cohen's standards) would equal the learning gain in one school year, which they consider an effect of huge practical relevance. Seen in this light the effect sizes that were found for a number of school effectiveness indicators (in particular school climate, curriculum quality, learning time and achievement orientation) should be upgraded in their rating for practical significance.



## ANNEX 3: More details on results from international comparative studies

In IEA studies and PISA, school, classroom and student level background variables from context questionnaires provide measures that can be associated with student performance. In most studies the school and student level context variables show a fair match with those addressed in school effectiveness research. This is of course a deliberate strategy, as one of the purposes of the international studies is to provide policy relevant explanations on performance differences between schools and countries, which is very similar to the “what works” mission of school effectiveness research. As an overarching re-analysis and overall review on “what works across countries”, based on these international assessment studies has not been carried out, to my knowledge, some miscellaneous study results will be briefly reviewed, before some tentative general trends will be formulated.

Bosker, (1997) carried out a secondary analyses of the IEA Reading Literacy Study, on which the basic report was published by Postlethwaite and Ross, in 1992. His results, when combining data from 27 countries, (100 schools per country) are summarized in the following citation (Bosker, 1997, in Scheerens and Bosker, 1997, p 254- 259):

“Both context indicators, public/private and rural/urban, show a positive association with adjusted school effects in reading; showing advantages for private and urban schools. From the input indicators, class size has a small, and meaningless, positive effect, and parental involvement has a clear positive effect (.08).

From the school process variables two achievement press variables (focus on higher order problem solving skills and focus on reading) have significant but small (.02) positive effects. The consensus & cooperation indicator has a signif-

icant but small (-.02) negative effect. The climate indicator shows a somewhat higher association (.04).

The other school process variables have estimated effects that are, statistically speaking, not discernable from zero. Of all teacher/classroom process variables only one has (an unexpected) negative effect: -.02, namely the effect of time for reading.

All in all the model for the international data does poorly, with only 9 percent of unique variation between schools accounted for by the educational effectiveness variables”.

In a secondary analysis of PISA 2000, focused at reading literacy performance, Luyten et al. (2005) looked at the impact of a number of malleable school variables, related to school resources, school climate and school policies. The impact of these school variables on reading literacy performance, expressed as the percentage of between school variance explained by these categories of variables was 2.1% for the school resources, 2.2 % for the school policies and 7.7 % for the climate variables. In comparison, the percentage of between school variance explained jointly by student socio economic background at individual and school level, was 51% (resources), 50.1% (policies) and 35.8% (climate). These authors also presented data on the number of countries in which specific school variables were associated statistically significant with reading literacy performance. The variable that reached statistical significance in relatively most countries was the index of disciplinary climate; this variable was significant in 11 of the 39 countries. Among the school resources variable “proportion of teachers with a third level qualification in language of assessment” had the highest number of statistically significant associations at country level, but this number was limited to just 4 countries. A similar number of countries (4) scored statistically significant on the variable “Students’ admission is considered for school admission”, as relatively the most frequent school policy variable reaching statistical significance at country level.

A final result in this study that sheds doubt on the impact of malleable school variables was the finding of the relatively large joint effect of the climate variables and school composition, expressed as school average socioeconomic status of the students. The OECD average for this component was 31.1% of the between school variation. In some countries this joint effect was high as 66.5 (Argentina) and 63.7 (Portugal). Scandinavian countries like Iceland (0.3%) and Finland (6/5%) had low joint effects.

Volume IV of the report on PISA 2009, (OECD, 2010) with reading literacy as the core performance domain is titled; "What makes a school successful; resources, policies and practices". The report provides information on the association of a number of system and school level variables with student achievement, more in particular results concerning selectivity and stratification (concepts of horizontal and vertical differentiation), decentralization and school autonomy, choice, accountability, educational resources (i.e. teacher salaries) and school climate.

The concept of vertical differentiation was coined to refer to the differentiation of 15-year old students across levels (grades) in educational systems, and to grade repetition. Horizontal differentiation refers to the degree of stratification of school structures, in terms of different school categories and age of first selection. The percentage of selective schools in the system was added as a third indicator of horizontal differentiation. The results indicate that strong vertical differentiation in terms of grade repetition is negatively associated with performance. In 24 OECD countries and 27 partner countries schools with more repeaters tended to achieve lower scores, after adjustment for student background. Grade repetition was also associated with low equity in the sense of high SES determinacy of performance.

Horizontal differentiation appeared to be unrelated to performance but tended to have a negative association with equity. School level indices of horizontal differen-

tiation, transfer of students to other schools because of low achievement and ability grouping appeared to be negatively associated with performance across countries. Country by country analyses showed a mixed pattern of positive and negative associations of these variables with performance. School systems with low levels of student transfer tended to have more school autonomy in the domain of curriculum and assessment.

School autonomy, free school choice and accountability arrangements are key features of school governance, addressed in volume IV. School autonomy in the domain of curriculum and assessment appeared to be positively associated with performance, while this positive association was not found with respect to autonomy in the use of school resources. Yet, in countries that had specific features of accountability in place, notably the posting of achievement data publicly by schools, school autonomy in the domain of resources did show a positive association with performance.

Across countries, school choice did not show a relationship with performance. Within countries, schools indicating a stronger level of competition tended to do better, but this association appeared to be highly dependent on the socioeconomic level of the school intake. This might be interpreted as a sign that high competition stimulates social segregation in schooling. Likewise the existence of private schools appeared to be unrelated to performance across countries, while within countries positive association depended mostly on the socioeconomic status of the students.

Standard based examinations appeared to be positively associated with performance across OECD countries. On average countries with standard based examinations scored 16 points higher on reading performance than countries that do not have these. No association was found for the use of standardized tests, across countries; within country analyses pointed out that several countries

did show a positive association. Use of assessment was positively associated with equity in the sense of relatively low SES determinacy of performance. Country by country analyses provided a somewhat mixed pattern of positive and negative associations as far as assessment is concerned.

Across countries, a positive association between educational resources, i.e. level of teacher salaries, and performance was found. Within school systems this association appeared to be strongly dependent on the socio economic background of the students. Provision of pre-school education had a positive association with equity. Resource related variables like instruction time, teacher/pupil ratio and expenditure did not show a positive association across OECD countries, but it did for all countries in PISA. The explanation that is suggested is that in industrialized countries resources do not dramatically vary between countries, and within countries, resources do not differ strongly between schools.

Of the variables that represented elements of the learning environment those associated with school climate (disciplinary climate and teacher student relationships) were most frequently associated with performance in country by country analyses. Other variables that showed some positive association were “positive behaviour among teachers”, and “parent pressure to achieve”. Across countries 3% of the student performance variation was attributable to the learning environment, when adjusting for student background; but jointly with SES the learning environment explained 9% of the variance. Even the most successful school variables had a significantly positive association with achievement in less than half of the participating countries. Typically, the texts of the OECD reports focus on the limited number of variables that show some effect, without mentioning the fact that a much larger set of variables did not show any significant associations.

Because all of the illustrative reviews of international

studies discussed in the above were about reading literacy performance, one further example will be given, in which science performance was the effect variable. It might be argued that reading literacy is less exclusively influenced by teaching at school than subjects like mathematics and science, and therefore might show lower effect sizes of malleable conditions at school and classroom level. In the report of the PISA 2006 edition (OECD, 2007), 5 school variables had small effects on student performance, after accounting for student background conditions: ability grouping (negative effect), high academic selectivity of school admittance (positive effect), schools’ posting achievement data publicly (positive effect), school average time students invested in learning science (positive effect) and school activities promoting students’ learning of science (positive effect), (OECD, 2007, p 267). These results are not dramatically more positive than those that were noted for reading literacy achievement; although PISA 2009, with reading literacy as the effect variable, showed a negative effect of school time (Scheerens, et al., 2013). The decomposition of total performance variation in science, as reported for PISA 2006, showed that 3.4% of the variance was explained by identified school and system level variables, net of student background conditions.

A few other secondary analyses looked at school leadership, time and opportunity to learn, respectively. Witziers et al. (2003) analysed data on school leadership in international comparative studies and found an average effect size of .02 (correlation). Baker et al. (2004) analysed data from three international comparative assessment studies: PISA, 2000, TIMSS, 1999 and the IEA CIVICS study (1999). Their conclusion about looking at the effect of time at between-country level is as follows: “As a number of studies have shown, we find that there is no significant relationship at the cross-national level between the achievement test scores and the amount of instructional time” (ibid 322). When they looked at effects of time within countries they found about an equal amount of countries

that had either a positive or a negative effect of time on achievement. Schmidt et al. (2001) found mostly significant correlations between curriculum sub-topics, (opportunity to learn) and achievement within nations.

What can we make of these results from international comparative assessment studies? The OECD is quite explicit on the limitations of the PISA data sets, as a basis for drawing conclusions about the influence of malleable system and school conditions. The sampling design of PISA is not given shape primarily to answer questions about school effects. Response on school variables is based on information from principals who are asked to generalize over the behaviour of many teachers. Student performance depends on previous years of schooling, beyond the time the students belong to the school where data collection takes place (".. the contextual data collected by PISA is an imperfect proxy for the cumulative learning environments of students, and their effect on learning outcomes is therefore likely to be underestimated" - OECD, 2007, p. 215). Finally, the study of school resources requires precision that might not be easily captured in the surveys (ibid).

Given these limitations international studies have difficulty in detecting school effects, which, even in research studies, show up as relatively small. In this way the results from international studies can be seen as a conservative test of "what works in schooling". Variables that appear to do the best in surviving this conservative test are: opportunity to learn (match between content covered and content that is tested), disciplinary climate, and use of evaluation and assessment for formative application as well as accountability purposes.

## ANNEX 4: Component and sub-items of fourteen effectiveness-enhancing factors (cited from Scheerens et al., 2007)

Factors	Components	Sub-components and exemplary items
1. Achievement, orientation, high expectations	1.1 clear focus on the mastering of basic subjects	The emphasis the principal places on the learning of basic skills like reading and science as a contrast to social and creative skills.
	1.2 high expectations (school level)	1) What percent of the students in this school do you expect to complete high school? 2) What percent of the students in this school do you expect to complete a 4-year college degree?
	1.3 high expectations (teacher level)	4 items measuring teacher's pro-achievement beliefs. Sample items: 1) The teachers in this school believe that learning is important. 2) The teachers at this school really believe that all pupils can achieve. 3) The teachers are only interested in the pupils who do well in tests and exams. 4) The teachers in this school seem to like teaching
	1.4 records on pupils' achievement	<ul style="list-style-type: none"> <li>the school keeps achievement records on all pupils</li> <li>the school uses achievement records to compare itself with other schools and with earlier performance</li> </ul>

2. Educational leadership	2.1 general leadership skills	5 items measuring firm and purposeful head teacher: 1) The head teacher takes action if a teacher's performance is inadequate. 2) The head teacher ensures that teachers are given support to improve their teaching, if they need it. 3) The head teacher encourages staff more than he/she criticises them. 4) The head teacher makes clear that the quality of teaching and learning at this school are his/her foremost priority
	2.2 school leader as information provider	Degree, timeliness and quality of information provision. The head teacher ensures that there is enough information on the work of colleagues in order to reach sufficient coordination of tasks.
	2.3 orchestrator of participative decision making	Degree to which principals and teachers shared equally in decision-making process. Sample item: how are decisions made at your school for: student retention policies, use of school funds; supplies and computers, selecting methods (only administrator decides....administrators and teachers decide jointly)
	2.4 school leader as coordinator	The school leader as an initiator of staff meetings
	2.5 meta-controller of classroom processes	9 items measuring integrated leadership: e.g., (a) There is evidence of significant instructional leadership in the school, (b) significant instructional leadership comes from a principal or other school-based administrator, and significant instructional leadership comes from teacher or group of teachers. (d) the actual influence of teachers over curriculum (e) the actual influence of teachers over instruction (f) the actual influence of teachers over student assessment (g) the actual influence of principals over curriculum (h) the actual influence of principals over instruction, and (i) the actual influence of principals over student assessment
	2.6 time educational/administrative leadership	<ul style="list-style-type: none"> <li>the number of hours a head teacher teaches</li> <li>total number of hours for managerial, non-teaching activities</li> <li>division of school leader activities over administrative/organizational, instructional leadership, contacts with parents, own professional development</li> <li>the number of times per year/month a head teacher attends lessons, discusses pupils' functioning with teachers</li> <li>teachers are content with the relative emphasis the head teacher spends on instructional versus other leadership tasks</li> <li>the degree to which teachers are satisfied with stimulating effectiveness enhancing leadership</li> </ul>
	2.7 counsellor and quality controller of classroom teachers	The concept is measured by three components: principal observation of classes (frequency p/y); evaluation of teachers (scale 1-3); evaluation of school quality by principal (scale 1-3)
	2.8 initiator and facilitator of staff professionalization	<ul style="list-style-type: none"> <li>the head teacher encourages further education of teachers in a selective, targeted way</li> </ul>



3. Consensus and cohesion among staff	3.1 types and frequency of meetings and consultations	<ul style="list-style-type: none"> <li>• Number of formal staff meetings with the head teacher</li> <li>• Frequency of informal meetings among groups of teachers</li> <li>• Informal contacts between staff</li> </ul>
	3.2 contents of cooperation	The degree of consistent practice among teaching staff indicated by within school standard deviation; one scale consisting of 8 items measuring the extent to which teachers' testing and grading practices are regulated (by their fellow teachers). Items concern assignment of report grades, time of announcing tests, content of tests, grading of tests, frequency of tests, discussing test results with class, frequency of quizzes. Five point scale (1=no rules- 5 formal rules exist and are influential: Items concern textbook, content to be covered, learning goals to be achieved, sequence of topics, amount of time spent on topics, homework, assignments to be made, teaching methods to be used. Five point Likert scale (1=no rules-5 formal rules exist and are influential)
	3.3 satisfaction about cooperation	<ul style="list-style-type: none"> <li>• satisfaction in relation to colleagues with respect to allocation of duties and coordination concerning: <ul style="list-style-type: none"> <li>- variety in interests</li> <li>- professional competence</li> <li>- supporting school improvement</li> <li>- involvement in pupils' learning and satisfaction</li> <li>- the amount of curriculum/'techniques'</li> <li>- discussion in team meetings</li> <li>- acceptance, support and opportunity to cooperate</li> <li>- cooperation at school and within the team</li> </ul> </li> </ul>
	3.4 importance attributed to cooperation	Measured by three scales: time for teacher collaboration (time), improving instruction through discussion (innovation), encouragement of encouragement in teacher participation (participate)
	3.5 indicators of successful cooperation	4 items asking the teachers and head teacher on the school's aims and values and how these are implemented through teaching and learning. Sample items: 1) Teachers at this school are all committed to the school's aims and values. 2) Teachers follow the same set of rules about pupil behaviour. 3) There is general agreement amongst the teachers about what are effective teaching. 4) Teachers and the head teacher agree on how teachers and pupils should behave towards each other. The degree to which teachers share similar goals and beliefs. Sample items: Most of colleagues share my beliefs and values about what the central mission of the school should be. Goals and priorities for the school are clear (strongly disagree....strongly agree)

4. Curriculum quality/opportunity to learn	4.1 the way curricular priorities are set	<p>Scale of 3 items measuring : the principal watches over the implementation of subject curricula (range 1-9)</p> <ul style="list-style-type: none"> <li>the extent to which subject matter provision is determined (i.e. guidelines are developed) by the ministry, the school board, the school team</li> <li>knowledge about core objectives arithmetic/math and science, the school work plan</li> <li>the importance of a good range of extra-curricular activities for the school's effectiveness</li> </ul>
	4.2 choice of methods and text books	<ul style="list-style-type: none"> <li>availability of books for language and math</li> <li>well-functioning methods for spelling, decoding, reading comprehension, composition writing and math, meaning: <ul style="list-style-type: none"> <li>a clear line with regard to subject matter content</li> <li>clear directives for instruction and testing</li> <li>a step-by-step approach for the low achievers</li> <li>a clear distribution of minimum competency goals over school years</li> </ul> </li> <li>which language methods (in which group)</li> <li>which arithmetic-math methods (in which group)</li> <li>method for science</li> </ul>
	4.3 application of methods and text books	<ul style="list-style-type: none"> <li>knowledge of the manual for arithmetic/math/science methods</li> <li>the time the method is being used</li> <li>considering transfer to other methods</li> <li>which part and which chapter in the beginning of the school year</li> <li>which part and which chapter now</li> <li>keeping sequence in the method</li> <li>% of subject matter dealt with at the end of the school year</li> <li>progress in method at the end of the school year</li> <li>other material for arithmetic/math/language/science than prescribed in method</li> <li>use of a calculator</li> <li>% of pupils being in a position to use a calculator</li> </ul>
	4.4 opportunity to learn	<ul style="list-style-type: none"> <li>% of time for arithmetic/math/science spent on method</li> <li>division of lessons to subject matter components</li> <li>other subject matter areas (within the subject)</li> <li>number of lessons per subject matter area</li> <li>which test items link up with education taught so far (for arithmetic/math and science)</li> </ul>

	4.5 satisfaction with the curriculum	<ul style="list-style-type: none"> <li>the extent of satisfaction with the curriculum now and 5 years ago</li> <li>satisfaction with the curriculum and the teaching materials</li> <li>satisfaction with the choice of subjects offered</li> <li>effectiveness of the curriculum's coordination within in the school</li> <li>successes with respect to extra-curricular activities and curriculum development over the past 5 years</li> <li>the degree to which the work at school is considering interesting</li> <li>the extent to which a curriculum is modern</li> <li>lessons: <ul style="list-style-type: none"> <li>number of lessons that stir the imagination</li> <li>diversity of subjects</li> </ul> </li> </ul>
5. School climate	a) orderly atmospheres	
	5.1 the importance given to an orderly climate	Scale of 7 items measuring : teachers are keen on a pretty and tidy classroom, students are encouraged to act orderly (range 1-21)
	5.2 rules and regulations	One scale consisting of 4 items measuring whether teacher made rules exist to control student behaviour. Items concern student being late, truancy, classroom disruption, cheating at tests. Five point scale (1=no rules - 5=formal rules exist and influence my work).
	5.3 punishment and rewarding	<ul style="list-style-type: none"> <li>% of pupils being disciplinary punished last year</li> <li>number of rewards mentioned by the school head</li> <li>number of punishments mentioned by the school head</li> <li>rewards/punishments ratio</li> <li>teacher rewards work more than punishment</li> <li>teacher rewards behaviour more than punishment</li> <li>forms of rewards by school head (a.o. praise)</li> <li>forms of punishments by school head (a.o. verbal warnings, confinement)</li> <li>a clearly applied system of punishment and rewarding at the school</li> </ul>
	5.4 absenteeism and drop out	School social problems scale includes 22 items that tap principal ratings of criminal activity, attendance problems, high-risk activities, and school-level parental investment in student health and wellbeing. (high scores = more problems).

5.5 good conduct and behaviour of pupils	<ul style="list-style-type: none"> <li>• other pupils do not encourage a child teasing another child</li> <li>• teachers and pupils see to it that teaching-learning processes are undisturbed</li> <li>• teachers create a learning environment in which pupils can work in a task-oriented way</li> <li>• see to it that nobody disturbs a teacher during the lesson</li> <li>• the pupils behave well when the teacher leaves the classroom</li> <li>• the lessons are not often disturbed by noise down the hall</li> <li>• level of pupil-sound in the classroom</li> <li>• level of pupil movement in the classroom</li> <li>• teachers' audibility in the classroom</li> <li>• pupils' behaviour around the school</li> <li>• strengthening pupils' behaviour</li> <li>• the level of unaccepted pupils' behaviour now and 5 years ago</li> <li>• important successes and problems with respect to pupils' behaviour and discipline now and 5 years ago</li> <li>• the school's high standards of pupil behaviour</li> <li>• the frequency school heads or team are being confronted with the following behaviour (of grade 6) <ul style="list-style-type: none"> <li>- vandalism</li> <li>- theft</li> </ul> </li> </ul>
5.6 satisfaction with orderly school climate	<ul style="list-style-type: none"> <li>• satisfaction with respect to safety at school, behaviour in the classroom, the school and teachers being attentive</li> <li>• satisfaction with respect to pupils' behaviour</li> <li>• degree of satisfaction with pupils' behaviour now and 5 years ago</li> <li>• the extent to which teachers set an example in their behaviour to pupils</li> <li>• satisfaction with respect to precautions/the way the school handles vandalism, drugs, alcohol and tobacco</li> </ul>
b) climate in terms of effectiveness orientation and good internal relationships	
5.7 priorities in an effectiveness-enhancing school climate	<ul style="list-style-type: none"> <li>• effectiveness enhancing conditions for a school <ul style="list-style-type: none"> <li>- a caring pastoral environment</li> <li>- positive inter-personal relationships for staff and students</li> <li>- the encouragement of a positive attitude to school (pride in school)</li> <li>- shared goals and values by staff and students</li> <li>- high level of pupil motivation</li> <li>- students satisfaction</li> </ul> </li> <li>• effectiveness enhancing conditions for your school <ul style="list-style-type: none"> <li>- students feel valued as people</li> <li>- encouragement of student responsibility</li> </ul> </li> </ul>

5.8 perceptions on effectiveness-enhancing conditions	<ul style="list-style-type: none"> <li>• effectiveness enhancing conditions of a school: <ul style="list-style-type: none"> <li>- teacher motivation</li> <li>- teacher commitment/effort</li> <li>- personal effectiveness of teaching staff</li> <li>- commitment/enthusiasm of teaching staff</li> </ul> </li> <li>• effectiveness restricting conditions of a school: <ul style="list-style-type: none"> <li>- heavy workload</li> <li>- low staff morale</li> <li>- lack of commitment and enthusiasm by some staff</li> <li>- high teaching staff absence rates</li> </ul> </li> </ul>
5.9 relationships between pupils	<ul style="list-style-type: none"> <li>• how do you feel about relationships between pupils</li> <li>• communication between pupils</li> <li>• pupils want to belong to the school and to each other</li> </ul>
5.10 relationships between teacher and pupils	<ul style="list-style-type: none"> <li>• how do you feel about relationships between pupils and teachers</li> <li>• contacts with pupils are open and pleasant</li> <li>• the teacher/pupil social relations are good</li> <li>• the team tries to understand pupils' needs</li> <li>• communication with teachers</li> <li>• teachers like pupils, support them, want them to associate nicely, know what every pupil wants, treat them fair, etc.</li> <li>• did the school have success with respect to better relationships between teachers and pupils the past 5 years</li> <li>• team functioning with respect to controlling pupils (firm but friendly relations)</li> </ul>
5.11 relationships between staff	4 items measuring warm staff atmosphere. Sample items: 1) The teachers at this school are friendly towards each other. 2) The teachers work well together. 3) If I have a problem, I will get support from other teachers. 4) Newly qualified teachers at this school are supported by experienced teachers
5.12 relationships: the role of the head teacher	<ul style="list-style-type: none"> <li>• communication between head teacher and pupils</li> <li>• head teacher listens to ideas/opinions/complaints from pupils about the climate and atmosphere)</li> <li>• relationships between school head and teachers</li> <li>• the school head: <ul style="list-style-type: none"> <li>- trusts his team members</li> <li>- can easily be approached</li> <li>- progresses job satisfaction</li> <li>- takes suggestions and ideas of teachers with respect to work climate and</li> <li>- sphere serious</li> <li>- pays attention to solving/improving mutual relations in case of conflicts</li> </ul> </li> <li>• the behaviour of school head evokes conflict</li> </ul>

5.13 engagement of pupils	<ul style="list-style-type: none"> <li>pupils have a say in what happens at school</li> <li>pupils co-decide about what happens at school</li> <li>pupils are proud of the school and show responsibility</li> <li>did the school have success with respect to pupils' responsibility the past 5 years</li> </ul>
5.14 appraisal of roles and tasks	<ul style="list-style-type: none"> <li>teaching/other tasks</li> <li>role clarity (clearly described tasks)</li> <li>job variety</li> <li>degree of job satisfaction</li> </ul>
5.15 job appraisal in terms of facilities, conditions of labour, task load and general satisfaction	<p>Job appraisal in terms of facilities, conditions of labour, task load and general satisfaction</p> <ul style="list-style-type: none"> <li>sufficient facilities (methods/materials) to efficiently carry out work</li> <li>salary and (secondary) conditions of labour</li> <li>competent authority passing on to a rewarding system based at personal commitment and motivation of teachers</li> <li>importance of part-time appointments</li> <li>opportunities for career enhancement</li> <li>task load (general anticipatory and perceived psychosocial mental strain): <ul style="list-style-type: none"> <li>- in general</li> <li>- own task load</li> </ul> </li> <li>satisfaction with respect to working-hours</li> <li>teachers believe they are overworked and under pressure</li> <li>average absenteeism of team members now and 5 years ago</li> <li>quality of working life</li> <li>satisfaction with respect to working with pupils</li> <li>enthusiasm for the work/the school (now and 5 years ago)</li> <li>attention for extra curricular activities</li> <li>feeling valued in functioning as a teacher</li> <li>opinion with respect to teachers' motivation</li> <li>successes/problems with respect to teachers' motivation during the past 5 years</li> </ul>
5.16 facilities and building	<ul style="list-style-type: none"> <li>classrooms/school/school building/playground clean, neat and well equipped</li> <li>sufficient space in/around the school</li> <li>sufficiently good facilities in and around the school</li> <li>no problems with respect to the school's entrance and with respect to stairs and halls in the school</li> <li>service quality in the area of safety, advice, care, health and canteen/stay-over facilities</li> </ul>



6. Evaluative potential	6.1 evaluation emphasis	<p>Measured whether or not there is a monitoring system, whether Ci-to-test is employed and whether there is a central registration of pupils' achievement.</p> <p>Scale of 9 items measuring "Evaluation policy" variable on School level: School wide use of tests for basic subjects, team evaluation of student progress, standardisation of achievement test procedures for basic subjects.</p>
	6.2 monitoring pupils' progress	<ul style="list-style-type: none"> <li>• a strong emphasis on the evaluation of test results</li> <li>• agreements and/or rules at school level with respect to testing/registration</li> <li>• at our school pupils' progress is regularly tested / we handle a good testing system for progress registration to register problems with pupils in time and to take appropriate measures</li> <li>• the extent to which a department head evaluates the learning progress in the department</li> <li>• in groups 1 and 2 attention is paid to early signaling so-called "pupils at risk" with regard to speech-language, social-emotional, auditive, visual-spatial and motor development, concern for more cognitive activities and the task and work attitude</li> <li>• the extent to which reading and arithmetic are tested</li> <li>• evaluation of pupils' progress takes place by means of standardized progress tests</li> <li>• what is pupils' assessment based on (national standards, comparison with other schools, progress of the child itself)</li> <li>• does the school handle achievement standards for individual pupils/standards at school level</li> <li>• (written) rules for promotion to the next year/retention yes/no</li> <li>• decision on promotion/retention based on opinion teacher</li> <li>• is the school posted on pupils' functioning in further education</li> </ul>
	6.3 use of pupil monitoring systems	<ul style="list-style-type: none"> <li>• pupils' progress being administered in a pupil monitoring system at school level</li> <li>• evaluating pupils' progress in basic skills at least twice a year by means of a pupil monitoring system</li> <li>• registration of pupils' progress in individual pupil files, in group surveys, in central pupil monitoring system</li> <li>• which pupil monitoring system is being used and do all teachers use the same pupil monitoring system</li> </ul>

	6.4 school process evaluation	<ul style="list-style-type: none"> <li>• has the school been assessed during the past 5 years by means of an instrument for school self evaluation</li> <li>• which aspects are structural tested/evaluated, analyzed and, if necessary, improved: <ul style="list-style-type: none"> <li>- pupil satisfaction</li> <li>- teacher achievement on the basis of pupil data</li> <li>- teacher satisfaction on the basis of .....</li> <li>- functioning of the school management</li> <li>- resource expenditure</li> <li>- courses and teaching</li> <li>- provision of education</li> <li>- new teaching methods</li> <li>- dissemination of innovations</li> <li>- the process of educational improvement</li> <li>- implemented changes</li> <li>- policy formation</li> </ul> </li> </ul>
	6.5 use of evaluation results	<ul style="list-style-type: none"> <li>• the school being aware of possible level of changes in pupil performance during the past 5 years</li> <li>• the school being aware of it's position with respect to pupil performance with regard to other schools having a comparable pupil population</li> <li>• for how many subjects is it possible to compare the present average achievement level to 5 years ago</li> <li>• for how many subjects does the school compare pupil progress with other schools</li> <li>• discussing pupils' progress and development regularly and systematically</li> <li>• evaluation of pupil performance: <ul style="list-style-type: none"> <li>- leads to adjustment of instruction and learning strategies</li> <li>- supports assignment to ability groups</li> <li>- changes in teaching strategies</li> </ul> </li> <li>• comparisons in achievement are being used for educational improvement</li> <li>• using former pupil data for educational improvement</li> </ul>

6.6 keeping records on pupils' performance	<ul style="list-style-type: none"> <li>• is keeping records on pupils' performance dealt with in the school work plan</li> <li>• if yes, indications for keeping records on pupils' performance concern the recording of it</li> <li>• teachers keep records on pupils' development and progress</li> <li>• does the teacher keep records on language progress</li> <li>• total number of registrations by teacher</li> <li>• how often does keeping records on individual pupil's progress in documents open to the school head occur</li> <li>• method of registration of learning progress:             <ol style="list-style-type: none"> <li>a. standardized data</li> <li>b. judgement by individual teacher</li> <li>c. both a and b</li> <li>d. there is no registration</li> </ol> </li> <li>• registration school progress:             <ul style="list-style-type: none"> <li>- not</li> <li>- in individual pupil file</li> <li>- in group summary</li> <li>- in central pupil monitoring system</li> </ul> </li> <li>• are pupils' data kept up with through the entire school career</li> <li>• if yes, by means of automatized computer system</li> <li>• frequency in which summaries of registration data are presented:             <ul style="list-style-type: none"> <li>- per pupil</li> <li>- per teacher</li> </ul> </li> <li>• group summaries of pupils' achievement are made</li> <li>• use summaries per pupil/teacher for ...</li> <li>• record results written assignment</li> <li>• record test results</li> <li>• execute an error analysis</li> <li>• process pupils' achievement in pupil monitoring system at school level</li> <li>• frequency of written reports to parents (per school year/group)</li> <li>• quality of reporting of pupils' progress (all-embracing, exploratory and valuable information on pupil's progress)</li> <li>• the school pays a lot of attention to reporting towards pupils and parents</li> <li>• written pupils' report when pupils pass to next school year</li> </ul>
6.7 satisfaction with evaluation activities	<ul style="list-style-type: none"> <li>• the degree of satisfaction with the student assessment/monitoring system now and 5 years ago</li> <li>• during the past 5 years, did the school succeed in establishing:             <ul style="list-style-type: none"> <li>- improved record-keeping/student profiles</li> <li>- improved monitoring of pupils' progress</li> </ul> </li> <li>• the team's satisfaction with respect to the amount of attention paid to improving education</li> </ul>

7. Parental involvement	7.1 emphasis on parental involvement in school policy	<ul style="list-style-type: none"> <li>• strong parental support as an important condition for school effectiveness</li> <li>• little parental support impedes effectiveness</li> <li>• school heads and teachers are open for suggestions from parents</li> <li>• the school emphasizing the importance of parental involvement with respect to education and pedagogical affairs</li> <li>• the school being open for parents attending lessons</li> <li>• the school has a parents' association of which parents can become a member on a voluntary base</li> <li>• are parents in parents' committees, parents' councils or participation councils reflecting the pupils' population and is this aimed for</li> <li>• agreements with respect to home visits</li> <li>• facilities for parents to be present in the school</li> <li>• parents' complaints are taken seriously</li> <li>• agreement with the following pronouncements: <ul style="list-style-type: none"> <li>- parental involvement is considered positive</li> <li>- parents are allowed to influence education's organizational structure</li> <li>- parents are allowed to influence educational contents</li> <li>- the school's and parents' responsibilities should be clearly defined</li> <li>- disappointing achievement is often due to parents not supporting the school</li> </ul> </li> <li>• a parent activity program is drafted yearly</li> <li>• the school stimulates that as many parents as possible attend the individual talks about their child's progress</li> <li>• the school pays specific attention to parents who are hard to reach</li> </ul>
	7.2 contacts with parents	"Parental Involvement" was created by averaging the responses to three items: whether the parent participates in school-related activities (1=seldom, 2=sometimes; 3=always); whether the parent knew his/her child's teacher meetings (1=never, 2=a little, 3=a lot), and whether the parent attended parent-teacher meetings (1=never or seldom, 2=almost always, 3=always).
	7.3 satisfaction with parental involvement	Parents' self-efficacy beliefs for their children's reading achievement were measured using an 18-item questionnaire by asking both parents perspectives, 10 items measuring parents' belief that they have the competence to successfully teach their children. sample item: 1) by reading to my child, I can help my child become a better reader. 8 items measuring parents' attributions for their children's success or failure. e.g., Children are good readers because they have a natural ability.

8. Classroom climate	8.1 relationships within the classroom	<p>Positive interaction with teacher: Likert scale based on frequency of 4 items; ranges from 0 (low) to 3 (high). Items: 1) Have you been told that your work is good? 2) Have you been asked questions in class? 3) Have you been praised for answering difficult question correctly? 4) Have you been praised because your written work is well done?</p> <p>Negative interaction with teacher: Likert scale based on frequency of 5 items; Values range from 0 (low) to 3 (high) Sample items: 1) Have you been given out to because your work is untidy or not done on time? 2) Have you wanted to ask or answer questions in class but were ignored? 3) Have you been given out to for misbehaving in class? 4) Teachers pay more attention in class to what some pupils say than to others. 5) I find most teachers hard to talk to.</p>
	8.2 order	<ul style="list-style-type: none"> <li>• fairness/firmness (control in the classroom)</li> <li>• classroom scores on: <ul style="list-style-type: none"> <li>- order in the classroom</li> </ul> </li> <li>• rules in the group are clear for each pupil</li> <li>• creation of an orderly, quiet work environment</li> <li>• situation with respect to control (firm but friendly relations) on pupils now and 5 years ago</li> </ul>
	8.3 work attitude	<p>Inattentive behaviours in the classroom were measured as dimensional, weighted composites of 4 items: 1) Cannot concentrate on any task; easily distracted 2) Lacks perseverance; is impatient with difficult or challenging tasks. 3) Easily frustrated; short attention span 4) Aimless; impulsive activity</p>
	8.4 satisfaction	<ul style="list-style-type: none"> <li>• Classroom fun factor The fun factor is to give an indication of whether or not it was an enjoyable experience to be a pupil in a particular teacher's class. the 'fun factor' is the sum of all 'yes' responses to the eight items that follow: <ul style="list-style-type: none"> <li>- Did the teacher smile often</li> <li>- Was there positive physical contact with pupils</li> <li>- Did the teacher show a sympathetic interest in the children other than as learners</li> <li>- Did the teacher chat to the pupils about non-work matters on any occasion during the day (Whether pupil or teacher initiated)</li> <li>- Was communication between children generally cheerful</li> <li>- Was the children's behaviour generally relaxed</li> <li>- Were there any jokes and/or was there any laughter in which the teacher was involved? (this does not include jokes at the expense of other pupils)</li> <li>- Was there any sign that pupils wanted to be in the classroom outside of class teaching time, either before or after sessions</li> </ul> </li> </ul>

9. Effective learning time	9.1 importance of effective learning	<ul style="list-style-type: none"> <li>• emphasis on               <ul style="list-style-type: none"> <li>- developing better policy and better procedures to enlarge instruction time</li> </ul> </li> <li>• impeding/progressing school effectiveness:               <ul style="list-style-type: none"> <li>- good registration of presence and absenteeism</li> <li>- good class management</li> <li>- give high priority to homework</li> </ul> </li> </ul>
	9.2 time	Scale of 6 items measuring: starting lessons on time, prevention of disturbances, rules on student truancy (range 1-18)
	9.3 monitoring of absenteeism	<ul style="list-style-type: none"> <li>• % of pupils truanting</li> <li>• the way the school handles absenteeism and lateness</li> <li>• satisfaction with respect to pupils' presence now and 5 years ago</li> </ul>
	9.4 time at school	<ul style="list-style-type: none"> <li>• number of school days</li> <li>• number of teaching days/hours               <ul style="list-style-type: none"> <li>- number of teaching days per school year</li> <li>- number of full teaching days per school week</li> <li>- number of semi teaching days per school week</li> <li>- total number of hours per school week</li> <li>- length of a school day</li> </ul> </li> <li>• % of cancelling of lessons</li> <li>• number of days with no lessons due to structural causes</li> <li>• % of total number of hours indicated on the table</li> <li>• measures to restrict cancelling of lessons as much as possible</li> <li>• policy with respect to unexpected absenteeism of a teacher</li> <li>• (in school work plan) agreements on substituting teachers</li> </ul>
	9.5 time at classroom level	<ul style="list-style-type: none"> <li>• number of lessons on timetable per school year</li> <li>• a lesson consists of how many minutes</li> <li>• amount of teaching hours for language/arithmetic</li> <li>• amount of minutes for arithmetic/physics per week</li> <li>• duration last arithmetic lesson in minutes</li> <li>• accuracy with respect to starting and finishing lessons in time now and 5 years ago</li> <li>• number of lessons that are cancelled</li> <li>• satisfaction with respect to available amount of time for working in the classroom</li> </ul>



	9.6 classroom management	<ul style="list-style-type: none"> <li>• attention for classroom management in the school work plan               <ul style="list-style-type: none"> <li>- with respect to lesson preparation</li> <li>- rules and procedures for the lesson's course</li> </ul> </li> <li>• situation with respect to aiming at work in the classroom (now and 5 years ago)</li> <li>• average % of teachers spending time on:               <ul style="list-style-type: none"> <li>- organization of the lesson</li> <li>- conversation (small talk)</li> <li>- interaction with respect to the work</li> <li>- supervision (pupil activities/behaviour)</li> <li>- feedback/acknowledgement</li> </ul> </li> <li>• average time during lesson spent on discussing homework, explaining new subject matter, maintaining order</li> <li>• sources of loss of time during lessons:               <ul style="list-style-type: none"> <li>- pupils do not know where to find equipment</li> <li>- disturbances due to bad behaviour of pupils</li> <li>- frequent interruptions</li> <li>- loss of time due to lengthy transitions from one activity to the next</li> <li>- unnecessary alterations in seating arrangements</li> <li>- frequent temporarily absence of pupils during lessons</li> <li>- waiting time for individual guidance</li> <li>- many (more than 3) teacher interventions to keep order</li> <li>- lack of control on pupils' task related work</li> </ul> </li> </ul>
	9.7 homework	<ul style="list-style-type: none"> <li>• attention for assigning homework at school/agreements in school work plan</li> <li>• homework after last (arithmetic) lesson: yes/no</li> <li>• number of homework assignments per week</li> <li>• type of homework (arithmetic/language) (reading/composition writing)</li> <li>• amount of homework</li> <li>• amount of time needed for homework (per day)</li> <li>• extra homework for low-achieving pupils</li> <li>• successes and problems now and 5 years ago with respect to:               <ul style="list-style-type: none"> <li>- prioritizing homework</li> <li>- a consistent homework policy</li> </ul> </li> <li>• whether homework assignments are graded or not.</li> </ul>

10. Structured instruction	10.1 importance of structured instruction	<ul style="list-style-type: none"> <li>• emphasis in school's policy on               <ul style="list-style-type: none"> <li>- the quality of teaching</li> <li>- encouraging pupils to take responsibility for their own learning process</li> <li>- (teacher independent learning)</li> <li>- emphasizing exam preparation</li> <li>- sufficient 'challenge' for both high and low achieving pupils</li> </ul> </li> <li>• to what extent agreed upon:               <ul style="list-style-type: none"> <li>- whole class instruction gives the best results</li> <li>- discovery learning mainly needs to happen outside the school</li> <li>- pupils acquire less knowledge when different pupils do different tasks</li> <li>- repeating a year often benefits pupils' development</li> <li>- the high-achieving pupil is especially the victim of individualized education</li> <li>- individualized education benefits all pupils</li> <li>- when dividing pupils into groups achievement will do as criterion</li> </ul> </li> </ul>
	10.2 structure of lessons	<ul style="list-style-type: none"> <li>• direct instruction divided in:               <ul style="list-style-type: none"> <li>- looking back daily</li> <li>- presenting subject matter</li> <li>- guided practice</li> <li>- giving feedback and correction</li> <li>- independent practice</li> <li>- looking back weekly/monthly</li> </ul> </li> <li>• teacher uses a lesson plan</li> </ul>
	10.3 preparation of lessons	<ul style="list-style-type: none"> <li>• lesson preparation building upon:               <ul style="list-style-type: none"> <li>- lessons formerly taught</li> <li>- written plan</li> <li>- other teachers/math specialists</li> <li>- text books</li> <li>- standardized tests</li> </ul> </li> <li>• most important information source for planning arithmetic/math lessons (lesson content, way of presentation, homework, tests)               <ul style="list-style-type: none"> <li>- core objectives</li> <li>- school work plan</li> <li>- manual</li> <li>- text book</li> <li>- other source books</li> </ul> </li> <li>• the subject matter is the central factor when teaching</li> </ul>

	10.4 direct instruction	<ul style="list-style-type: none"> <li>• attention for instruction in the school work plan</li> <li>• indications in school work plan with respect to: <ul style="list-style-type: none"> <li>- clear objectives of instruction</li> <li>- construction of the instruction</li> <li>- way of presenting subject matter</li> <li>- the use of instructional materials</li> </ul> </li> <li>• explanation or help to individual/groups of pupils in or outside the lesson</li> <li>• teachers deal with subject matter that corresponds to the lesson's aim</li> <li>• teacher explains at the beginning of the lesson to what prior knowledge the subject matter corresponds</li> <li>• teacher gives pupils the chance to raise questions about the last lesson</li> <li>• teacher explains beforehand what pupils have to know at the end of the lesson</li> <li>• teacher knows what to achieve with the lesson</li> <li>• lesson objectives are clear to pupils</li> <li>• teacher applies instructional methods to increase pupil's achievement</li> <li>• teacher deals with only one subject matter component at the time</li> <li>• explanation in small successive steps</li> <li>• teacher takes next step when preceding step is understood</li> <li>• teacher gives concrete examples</li> <li>• it appears from pupils' reactions that the teacher explains the subject matter clearly</li> <li>• teacher poses intellectual questions that invite pupils to participate actively</li> <li>• after posing a question the teacher waits to let the pupils think</li> <li>• teacher gives many pupils a turn</li> <li>• a lot of interaction between teacher and pupils</li> <li>• pupils respond well to questions posed by the teacher</li> <li>• teacher have pupils practised under guidance</li> <li>• teacher continues until all pupils have mastered the subject matter</li> <li>• explanation is clear</li> <li>• teacher involves pupils in instruction</li> <li>• teacher takes care that pupils are concentrated during instruction</li> <li>• during instruction immediate feedback to answers of pupils</li> <li>• the lesson displays a clear structure</li> <li>• at the end of instruction summary of subject matter (by teacher/pupils)</li> <li>• pupils get tasks they can handle</li> <li>• group work, if appropriate</li> <li>• teacher's activities (controlling) when pupils work on assignments</li> </ul>
--	-------------------------	--

		<ul style="list-style-type: none"> <li>• teachers take time to help pupils with tasks</li> <li>• pupils know which tasks are to be carried out</li> <li>• teacher sees to it that pupils work in a concentrated way during assignments</li> <li>• teacher sees to it that pupils work task-oriented during assignments</li> <li>• from pupils' reactions it appears that everyone knows what he or she has to do</li> <li>• there is sufficient control on pupils doing the assignments they are supposed to do</li> <li>• pupils work at a good pace</li> <li>• % of time during lessons in which assignments are discussed</li> <li>• analysis of mistakes</li> <li>• checks on homework</li> </ul>
	10.5 monitoring	<ul style="list-style-type: none"> <li>• is monitoring of pupils' achievement mentioned in the school work plan</li> <li>• indications concerning: <ul style="list-style-type: none"> <li>- pupils' written assignment</li> <li>- the use of tests</li> </ul> </li> <li>• % of lessons containing tests</li> <li>• the number of tests, hearings</li> <li>• types of tests per school year (a.o. posing questions in class, own tests, curriculum-embedded tests)</li> <li>• which procedures are used to assess pupils' achievement with respect to arithmetic</li> <li>• progress in pupil learning outcomes is measured by means of (curriculum-embedded) tests</li> <li>• teacher uses checklist for oral hearing of pupils</li> <li>• the way the teacher prepares pupils for tests</li> <li>• teacher checks whether all pupils have reached the minimum goals</li> <li>• teacher checks up on difference between expected and actual pupil achievement</li> <li>• compare pupil achievement to: <ul style="list-style-type: none"> <li>- former pupil achievement</li> <li>- fellow-pupil achievement</li> <li>- norms and standards</li> </ul> </li> <li>• in what way is arithmetic/math work of a pupil judged (absolute criterion, class average etc.)</li> <li>• are test results used for individual help, extra explanation</li> <li>• taking action in connection with test results</li> </ul>

		<ul style="list-style-type: none"> <li>• use learning progress for: <ul style="list-style-type: none"> <li>- preparing a program for individual pupil</li> <li>- reporting to parents</li> <li>- informing teacher about next group</li> <li>- evaluating the school's functioning</li> <li>- putting pupils into (parallel) classes</li> <li>- selecting pupils for teaching programs (enrichment/remediation)</li> <li>- grouping pupils within classes</li> <li>- other</li> </ul> </li> <li>• the degree of pupils' progress has an effect on class level (e.g. other grouping patterns, more or less instruction etc.)</li> <li>• successes/problems with respect to preparation for tests over the past 5 years</li> <li>• review and correct written assignment of pupils</li> <li>• use of curriculum-embedded tests</li> <li>• use of curriculum-independent tests</li> <li>• use of self-made tests</li> </ul>
11. Independent learning	no sub-components	<ul style="list-style-type: none"> <li>• attention for independent learning in school work plan</li> <li>• teacher-independent learning is being encouraged yes/no</li> <li>• if yes, indications concerning: <ul style="list-style-type: none"> <li>- relation instruction/processing time</li> <li>- organization of independent learning</li> <li>- other types of differentiation</li> </ul> </li> <li>• state of affairs with respect to teacher-independent learning/independent learning</li> <li>• the extent to which pupils are responsible for their own work</li> <li>• the extent to which pupils are responsible for their own work during a longer period</li> <li>• the extent to which pupils are able to chose their own assignments</li> <li>• the extent to which pupils' cooperation is encouraged by teachers</li> <li>• in case of independent learning, do pupils work: <ul style="list-style-type: none"> <li>- on the same subject</li> <li>- on various subjects per group of same level</li> <li>- on the same subject at own level</li> <li>- on various subjects at various levels</li> </ul> </li> <li>• opportunity for pupils to plan the school day themselves</li> <li>• successes and problems with respect to teacher-independent learning/independent learning</li> </ul>

12. Differentiation	12.1 general orientation	<p>The extent of streaming and associated curricular differentiation in the school; Guttman scale ranging from 0 (mixed ability base classes) to 4 (highly streamed)</p> <ul style="list-style-type: none"> <li>• how to deal with differences between pupils in arithmetic/math attainment levels during lessons (all pupils the same subject matter, ...)</li> <li>• % of lessons in which pupils: <ul style="list-style-type: none"> <li>- work on the same subject</li> <li>- work on two subjects</li> <li>- work on three or more subjects</li> </ul> </li> <li>• how often do pupils work individually or in pairs</li> <li>• % of teacher time spent on communication with the class, groups and individuals</li> <li>• criteria with respect to subject matter provision/grouping: <ul style="list-style-type: none"> <li>- achievement</li> <li>- results standardized test</li> <li>- results diagnostic test</li> <li>- results oral test</li> <li>- teachers' recommendations</li> <li>- parents' wishes</li> <li>- pupils' wishes</li> <li>- method's demands</li> </ul> </li> <li>• pupil grouping within the class: <ul style="list-style-type: none"> <li>- no grouping</li> <li>- age groups</li> <li>- level groups</li> <li>- interest groups</li> <li>- other</li> </ul> </li> <li>• frequency of regrouping pupils (evt. of more classes) on behalf of level groups</li> <li>• problems and successes with respect to differentiation the past 5 years</li> <li>• subject matter mastery adapted to slow and fast learners</li> </ul>
---------------------	--------------------------	---



	12.2 special attention for pupils at risk	<ul style="list-style-type: none"> <li>• policy with regard to low-achieving pupils</li> <li>• school policy is explicitly aimed at catering for a wide range of educational needs: in other words, clear directives and structural attention for pupils with problems</li> <li>• catering for special individual educational needs concerning: <ul style="list-style-type: none"> <li>- diagnosing pupils "at risk"</li> <li>- remedial teaching</li> <li>- cooperation with special education</li> <li>- drafting intervention plans</li> <li>- drafting group plans</li> </ul> </li> <li>• amount of extra time teachers are prepared to spend on problem pupils</li> <li>• extra provisions for problem pupils</li> <li>• low-achieving pupils get more time for reflection, extra attention, instruction, help, material and exercise material</li> <li>• provisions/approved methods for preventing (teaching) problems</li> <li>• check systematically which subject matter is not being mastered</li> <li>• group teachers having expertise with regard to diagnostic test administration</li> <li>• group teachers are able to translate test data into intervention plans</li> </ul>
13.	no sub-components	<p>Reinforcement</p> <ul style="list-style-type: none"> <li>• is feedback in connection with pupils' achievement discussed in the school work plan</li> <li>• indications for feedback in connection with pupils' achievement are related to discussion by the teacher</li> <li>• how often, in arithmetic/math lessons, do you take the following action when pupils answer wrongly (a.o. correct wrong answer, pose different question)</li> <li>• during the lesson feedback is given and pupils' mistakes are corrected</li> <li>• when pupils carried out an assignment it is discussed immediately</li> <li>• the teacher explains what was wrong when he returns the tests</li> <li>• teacher gives pupil as much as possible real and positive feedback to achieved results</li> <li>• frequency of discussing learning progress with pupils</li> <li>• low-achieving pupils get extra feedback</li> </ul>

		<p>Feedback</p> <ul style="list-style-type: none"> <li>• results written assignment is discussed with pupil if necessary</li> <li>• results curriculum-embedded test are discussed with pupil if necessary</li> <li>• results method-independent tests are discussed with pupil if necessary</li> <li>• results of self-made tests are discussed with pupil if necessary</li> <li>• a differentiated supply based on tests is offered</li> <li>• quality/suitability of feedback</li> <li>• state of affairs with respect to giving constructive feedback now and 5 years ago</li> <li>• problems with respect to inadequate feedback</li> </ul> <p>2 items measuring positive reinforcement (teachers): 1) The school has a system for rewarding pupils who work hard and/or make good progress even if they do not get high grades. 2) A pupil who works hard or makes good progress is noticed and praised.</p> <p>4 items measuring monitoring and rewards: 1) I am set targets for my learning by my teachers which are individual to me and not for the whole class. 2) The school has rewards for pupils who work hard or make good progress even if they do not get high grades. 3) A pupil who works hard or makes good progress is noticed and praised. 4) Teachers notice those pupils who are not working as well as they could and try to make them work harder.</p> <p>Measures the quality of teaching. Example: the frequency of monitoring whether pupils have mastered the learning content, the frequency of repeating learning content where necessary, the frequency of evaluating pupils' progress and giving help to pupils as needed.</p> <p>Measuring the degree of regular feedback given to pupils on achievement tests</p> <p>Scale measuring teacher feedback. Items use from the School Organisational Health Questionnaire (SOHQ) (teachers perceptions)</p>
--	--	---

## ANNEX 5: Annotated bibliography of key publications

Creemers, B.P.M., & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. London and New York: Routledge

The book describes a comprehensive multi-level model of educational effectiveness (system level, school level, classroom level and individual student level. It is called dynamic because longitudinal research is emphasized. An interesting phenomenon of the model is that key independent variables are measured according to various facets, like quality and frequency. The message is that factors at classroom level make most of the difference in educational effectiveness. This claim is supported on the basis of empirical studies and a meta-analysis.

Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.

Results of no less than 800 meta-analyses of educational effect studies are synthesized. Different sections deal with school level, teaching level and curriculum variables. Most variables indicate educationally significant effect sizes. According to the author the problem is not that “nothing works” in education, but rather that “everything works”. The author draws the evidence together by contrasting active, structured teaching approaches, to more “laissez faire” approaches, and provides evidence that the former is more effective.

Woessmann, L., Luedemann, E., Schuetz, G. & West, M.R. (2009). *School Accountability, Autonomy and Choice around the World*. Cheltenham, UK/ Northampton, MA, USA: Edward Elgar

Contains empirical results on the effect of system level characteristics of educational systems on student achievement. The results are based on international comparative assessment studies, TIMSS and PISA. Key variables that are addressed are decentralization and school auto-

my, accountability and examinations and school choice. Results of the effectiveness of these system level variables is rather mixed; specifically interesting are interaction effects. For example, central examinations appear to make more of a difference in systems with relatively large school autonomy. Another interesting feature is that effects are expressed in terms of quality (achievement levels) and equity; the degree to which attainment is conditional on the socio-economic status of the students.

Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S., (2003) Comprehensive school reform and achievement: a meta-analysis. *Review of Educational Research*. 73, 125-230

In this review- article a comprehensive overview is provided of Comprehensive School Reform, as an evidence based, multi-faceted approach to school improvement. Typically comprehensive school reform programs combine and integrate various effectiveness enhancing factors, at the level of school governance, school management, teacher professional development, structured teaching, and application of formative assessment and feedback. One of the internationally best known examples is the Success for All program, developed by Bob Slavin. The meta-analysis shows that the programs have positive effects, though of a fairly modest size.

Thomas, S.,M., Peng, W.,J., & Gray, J., (2010) Modeling patterns of improvement over time: value added trends in English secondary schools. *Oxford Review of Education*. 33, 261-295.

Thomas et al. (2010) analyzed school data over a period of 11 years in the UK Lancashire district. They concluded that there was a fair stability in school effects. Still, when schools were categorized as average, over- or underachieving there were many switches, and over a period of 11 years 50% of the schools had changed category. Moreover continuous progress was rare.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14.

The term pedagogical content knowledge “gained renewed emphasis with Lee Shulman (1986), a teacher education researcher who was interested in expanding and improving knowledge on teaching and teacher preparation that, in his view, ignored questions dealing with the content of the lessons taught. He argued that developing general pedagogical skills was insufficient for preparing content teachers as was education that stressed only content knowledge. In his view, the key to distinguishing the knowledge base of teaching rested at the intersection of content and pedagogy (Shulman, 1986)”. - See more at: [http://www.idra.org/IDRA\\_Newsletter/August\\_2009\\_Actionable\\_Knowledge/Pedagogical\\_Content\\_Knowledge/#sthash.UeB576U1.dpuf](http://www.idra.org/IDRA_Newsletter/August_2009_Actionable_Knowledge/Pedagogical_Content_Knowledge/#sthash.UeB576U1.dpuf)

Heck, R. H. & Moriyama, K.(2010). Examining relationships among elementary schools’ contexts, leadership, instructional practices, and added-year outcomes: a regression discontinuity approach. *School Effectiveness and School Improvement*, 21(4), 377-408.

This is an exemplary empirical study that illustrated recent developments of school leadership effects research. The study has a few strong assets: it uses an indirect effects model in a longitudinal design and successfully shows how leadership behaviour can influence instructional practices, which in their turn are influencing student achievement.

Scheerens, J., Luyten, H., & Van Ravens, J. (2011) *Perspectives on Educational Quality. Illustrative Outcomes on Primary and Secondary Education in the Netherlands*. Springer Briefs in Education, Dordrecht, Heidelberg, London, New York: Springer

The book starts out with a conceptualization of educational effectiveness as a facet of the broader concept of educational quality. Other facets are productivity, efficiency, equity and responsiveness. Various alternative interpretations of educational quality are discussed. The conceptual part is followed up with a more operational section on indicators of effectiveness enhancing conditions at system, school and classroom level. The second part of the study gives an analysis of educational achievement outcomes (test results) and attainment outcomes (realized school careers) in the Netherlands. The final discussion includes a concise overview of the current state of the art in educational effectiveness research.

Carnoy, M., Elmore, R., & Siskin, L., S. (2003) *The new accountability. High schools and high stakes testing*. New York: Routledge

This study is based on the comparative effectiveness of accountability policies across different states of the United States of America. One striking outcome is that states that employ more rigorous high stakes accountability policies tend to have better achievement outcomes. A second very interesting analysis relates “external” accountability, to what is described as “internal” accountability, which is described as the way in which schools are capable to use internal information and data for evaluation and planning purposes. It is suggested that the success of external accountability depends on the capacity of schools with respect to their “internal” accountability.

Slavin, R.E. (1998). *Sands, Bricks, and Seeds: School Change Strategies and Readiness for Reform*. Hargreaves, A., Lieberman, A., Fullan, M., And Hopkins, D. *International Handbook of Educational Change*.: Dordrecht, Boston, London: Kluwer

This chapter makes a case for a differentiated approach to school improvement, taking into consideration differ-

ent levels at which schools are already functioning.

Moreover the chapter builds a strong case for external support and intervention, particularly for low and average achieving schools. It illustrates a powerful alternative to the predominant mode of “bottom up” reform in the field of school improvement.